Journal of Knowledge Economy Studies

Vol. 3, No. 1
Spring 2026

Publisher:
Hazrat-e Masoumeh University

Director-in-Charge:
Marziyeh Nourahmadi

Editor-in-Chief:
Mona Jami Pour

Assistant Editor:
Marziyeh Nourahmadi

Journal Staff:
Fatemeh Ostadjafari

Language Editor:
Saeedeh Kavoshian

Technical Editor:
Zahra Ostovari

Online ISSN: 3060-7329

Knowledge Economy Studies is an open access double-blind peer reviewed publication which is published by Hazrat-e Masoumeh University. This journal is a quarterly publication, which publishes original research papers on journal scope. This journal follows Committee on Publication Ethics (COPE) and publishes research findings in fields related to the relationship between knowledge, technology and digital economy. All submitted manuscripts are checked for similarity through Samim Noor software to ensure their authenticity and then rigorously peer-reviewed by expert reviewers.

Address: Qom, Ghadir Boulvard, Opposite to Yadgar Imam Stadium, Hazrat-e Masoumeh University.

Tel: +9825 – 33209030
Email: jkes@hmu.ac.ir
Website: http://kes.hmu.ac.ir

Auhors Guidelines

To ensure geographical and organizational diversity among authors, each researcher will be allowed to publish only one article as the first or corresponding author. There is no limit to the number of articles that can be published as co-authors.

Formatting Guidelines

Please adhere to the following formatting guidelines for each section of the article. Ensure consistency in font type, size, and style as specified to maintain uniformity throughout the document.

Section	Font Style	Font Style	Font Size
Article Title	Bold	Cambria	14
Abstract Title	Bold	Cambria	12
Abstract	Regular	Cambria	9
Keywords Title	Bold	Cambria	11
Keywords	Regular	Cambria	9
Tables/Figure Title	Bold	Cambria	8.5
Text within tables/charts	Regular	Cambria	9
Main Headings	Bold	Cambria	12.5
Other Headings	Bold	Cambria	11.5
Text	Regular	Cambria	12
Header	Regular	Cambria	9
Footnotes	Regular	Cambria	9
References	Regular	Cambria	11

Content and Length of Manuscripts

Content and Length of Manuscripts the Editor welcomes original articles that fall within the aims and scope of the Journal and are as concise as the subject matter and research method permit. Manuscripts should be in English and, where possible, the text should be around 6000 to 8000 words. The first page of the text should begin with the title only (without the author's name) and an abstract of no more than 250 words. Include a list of up to five keywords suitable for indexing and abstracting services. This abstract should summarize the entire paper, not just the conclusions. Manuscripts must be typed with 1.15 line spacing throughout the main text. Submission of a paper or research note implies it contains original work that has not been published previously and is not under consideration for publication elsewhere. Copyright of all accepted contributions to the Journal will be vested in the University of Kurdistan.

Mode of Submission

All manuscripts should be submitted electronically via the KES online editorial office, Manuscript Central, at: https://kes.hmu.ac.ir.

Other Points

The originality of the article will be checked using similarity detection software. The maximum acceptable similarity percentage in such software is 15%. Authors should refrain from excessive self-citation in the submitted article. The entire or part of the article should not have been previously published in books, journals, or conferences. Except in special cases, referencing secondary sources should be avoided. Registering an ORCID code is mandatory for all authors. To do this, authors must visit https://orcid.org, register, and then enter their ORCID code in the author details section of the journal system. Author details, including name, surname, academic rank/institution, department name, institution name, city, country, email address, ORCID identifier, telephone number, and address, should be included on a separate page. Therefore, the article is submitted anonymously in a WORD file format.

References

References must be formatted with 1.0 line spacing to ensure clarity and efficient use of space.

Journal of Knowledge Economy Studies, Vol. 3, No. 1, Spring 2026

Contents

Journal of
Knowledge Economy Studies

https://kes.hmu.ac.ir

Hazrat-e Masoumeh University

Developing a Deep Learning–Based Model for Predicting and Detecting Fraud in Financial Statements

Article Type:
Research Article

Narges Mehrabi Hashtchin[1]
Faculty of Social Sciences and Economics, Alzahra University, Tehran, Iran
E-mail: n.mehrabi@alzahra.ac.ir

Gholamreza Soleymani Amiri[2*]
Faculty of Social Sciences and Economics, Alzahra University, Tehran, Iran
Corresponding Author
E-mail: gh.soleymani@alzahra.ac.ir

Received: 27 January 2026
Revised: 17 February 2026
Accepted: 17 March 2026
Available Online: 18 March 2026

Abstract

This study developed a data-driven framework for financial statement fraud detection by benchmarking machine learning, deep learning, and hybrid classifiers under a unified, leakage-resistant evaluation protocol. The fraud cases were identified from the U.S. Securities and Exchange Commission's Accounting and Auditing Enforcement Releases (AAERs) and matched with Compustat data over 1991–2014, producing 122,526 firm-year observations, including 902 confirmed fraud cases. Four structured-input configurations were evaluated: 28 raw financial statement items, 14 financial ratios, their combined set (28+14), and a parsimonious seven-feature subset (six ratios plus Altman's Z-score). The features were selected using minimum redundancy–maximum relevance (mRMR), class imbalance was addressed via cost-sensitive learning, and performance was assessed with a firm-level 80/20 split and stratified group-based five-fold cross-validation within training. The empirical results indicated that deep and hybrid models consistently outperform classical tabular baselines, reflecting non-linear and interaction-driven fraud signals. The Transformer achieved the most stable and highest overall performance, reaching 0.98898 accuracy and a 0.51087 F1-score under the seven-feature configuration. The combined raw-item and ratio inputs outperformed the ratios alone, implying incremental predictive value in raw accounting items, while the best overall outcomes were obtained with parsimonious seven-feature subset. Collectively, the findings supported the study's hypotheses and demonstrated the effectiveness of attention-based modeling for financial statement fraud detection.

Keywords

Artificial intelligence, Fraudulent financial statements, Machine learning, Deep learning, Imbalanced datasets.

Cite this article: Mehrabi Hashtchin, N. & Soleymani Amiri, Gh., (2026). Developing a deep learning–based model for predicting and detecting fraud in financial statements. *Journal of Knowledge Economy Studies (JKES)*, 3(1), 7-24. DOI: http//doi.org/10.22034/kes.2026.2083828.1098

Online ISSN: 3060-7329

Introduction

Financial statement fraud remains a persistent concern for capital markets and financial oversight because intentional misreporting can distort valuation, impair resource allocation, and erode confidence in audited reporting. In practice, fraud detection is difficult because fraudulent behavior is typically adaptive, multi-faceted, and intentionally concealed within otherwise legitimate reporting processes, while audit procedures operate under time and evidence constraints. Consequently, an active research agenda has emerged around data-driven fraud-risk screening systems that can help auditors and regulators prioritize firm-year investigations and allocate attention to the highest-risk cases.

Recent research in fraud analytics has advanced along three complementary directions: (1) improved learning architectures for complex, non-linear patterns, (2) richer representations and modalities beyond conventional tabular inputs, and (3) increased emphasis on explainability and operational deployability. First, a growing body of evidence supported using modern machine learning ensembles and meta-classifiers for financial fraud prediction, where combining heterogeneous learners can yield stronger and more stable performance than individual models (Achakzai & Juan, 2022; Azim Mim et al., 2024; Cao et al., 2023).

This perspective is consistent with broader fraud-detection works showing that ensemble strategies, including soft-voting and boosting variants, can improve robustness under noisy signals and rare-event settings (Ahmed et al., 2025; Azim Mim et al., 2024).

Second, the scope of information sources used for fraud detection has broadened beyond the handcrafted ratio sets. In financial reporting contexts, recent studies demonstrated that unstructured disclosures contain incremental predictive content. Contextual language learning approaches based on Transformer-style architectures (e.g., BERT) can extract deception-related signals from narrative sections of annual reports, thereby improving accounting fraud detection performance and increasing the yield of detected fraudulent observations under limited investigative capacity (Bhattacharya & Mickovic, 2024). In parallel, work that constructs multi-dimensional firm portraits by combining structured indicators with broader firm characteristics and unstructured data further supports the view that richer representations can improve the fraud prediction and, importantly, can be paired with post-hoc explanation methods to identify salient risk drivers (Zhang et al., 2025). These developments collectively suggest that fraud-risk systems should be evaluated with attention to both predictive performance and the evidentiary value of the underlying signals.

Third, explainability has become a central requirement in high-stakes financial applications. Several recent contributions explicitly argued that many prior fraud-detection models remain difficult to interpret, which limits their acceptance in audit and regulatory environments. To address this, explainable frameworks have been proposed that embed accounting-relevant structure into the modeling pipeline. For instance, a two-layer knowledge-graph approach for financial statement fraud detection models semantic and articulation relationships among statement items to enable interpretable pattern mining and credible fraud assertions (Cai & Xie, 2024). Complementary research proposed interpretable graph-learning formulations that jointly address severe imbalance and provide built-in interpretability outputs, reflecting a broader movement toward accountable and deployable fraud analytics (Lu et al., 2026). Related work on graph-based fraud detection further indicated that relational structure and dependency modeling can improve detection capability in fraud settings characterized by complex interactions (Shao et al., 2026).

Across these lines of research, two technical challenges recur. The first is extreme class imbalance. Fraud is typically rare, so high overall accuracy can coexist with limited minority-class detection quality. Prior work in financial misstatement settings highlighted the importance of cost-sensitive learning and imbalance-aware modeling to reduce the risk of accuracy inflation and to improve performance on fraud-related outcomes (Kim et al., 2016; Lu et al., 2026). The second challenge is operational realism. Fraud tactics evolve and detection systems may need to function under limited labels, streaming data and scale constraints. Studies on transaction fraud increasingly emphasized online or adaptive frameworks, distributed learning, unsupervised anomaly detection, and hybrid sampling to improve practicality under real-world conditions (Ahmed et al., 2025; Karnavou et al., 2025; Lei et al., 2023; Narayana Gorle & Panigrahi, 2026). Although these works often targeted transaction-level fraud, they reinforced methodological expectations that are increasingly relevant for financial statement fraud. Stable generalization, careful evaluation under realistic splits, and transparent reasoning or explanation mechanisms were given due attention.

Despite substantial progress, an open and practically relevant question remains in financial statement fraud research regarding which structured inputs are most informative and deployable: raw financial statement items, derived financial ratios, or their combination. In addition, the literature increasingly calls for solutions that not only achieve strong predictive discrimination but also remain usable for practitioners through parsimonious feature sets and interpretable outputs (Cai & Xie, 2024; Zhang et al., 2025). Motivated by these directions, this study developed and evaluated a data-driven framework for financial statement fraud detection that systematically compares classical machine learning methods with modern deep and hybrid architectures under firm-level splitting and stratified group-based cross-validation. The researchers examined four input scenarios including raw financial variables, financial ratios, their combination, and a compact seven-feature subset to assess the incremental value of raw versus ratio-based information and to identify a parsimonious configuration suitable for fraud-risk monitoring and audit planning. To address the severe rarity of fraud cases, we incorporated imbalance-aware learning through cost-sensitive training, and we emphasized stability across validation folds to support robust model comparisons in applied settings (Kim et al., 2016; Lu et al., 2026).

Literature Review

Financial statement fraud is a major topic in accounting and auditing and refers to deliberate and deceptive actions undertaken to obtain benefits (Arboleda et al., 2018). Fraud in financial reporting refers to the misstatement of financial reports and the presentation of a distorted picture of the business entity (Vakilifard et al., 2009).

The aim of this study was to apply deep learning to predict and detect fraud in financial statements based on financial ratios and to compare its effectiveness with raw financial statement data. The fraud cases were extracted from material misstatements reported in the U.S. Securities and Exchange Commission's AAERs, because these releases rely on publicly available, low-cost data and enable fair comparison with prior studies. Drawing on the literature (Ahmed & Courtis, 2015; Kanapickienė & Grundienė, 2015), financial ratios due to their grounding in accounting expertise, simplicity, and widespread use can serve as more effective indicators for fraud detection. This study employed the validated dataset of Bao et al.'s study (2020), which includes both raw features and financial ratios, and by comparing these inputs, identified the most effective feature set for improving the fraud detection accuracy.

Table 1 synthesizes representative prior studies and highlights heterogeneity in data domains, input representations, model families, and evaluation protocols, which complicates direct cross-study inference.

Table 1.
Representative Studies and Design Choices in Fraud Detection Research

Study	Data domain	Inputs	Model family	Validation design
Dechow et al. (2011)	Financial statements	Ratios / engineered	Benchmark statistical	Holdout
Cecchini et al. (2010)	Financial statements	Raw items	Classical ML (SVM)	Holdout
Bao et al. (2020)	Financial statements	Raw items (+ ratio benchmarks)	ML / ensemble	Holdout
Bhattacharya & Mickovic (2024)	Disclosures	Text sequences	Transformer/NLP (BERT)	Holdout
Cai & Xie (2024)	Financial statements	Graph	Explainable KG + pattern mining	Benchmark evaluation
Zhang et al. (2025)	Financial statements	Structured + broader signals	ML + explainability (SHAP)	Holdout
Lu et al. (2026)	Financial fraud	Graph-structured features	Interpretable, imbalance-aware graph learning	Benchmark evaluation
Karnavou et al. (2025)	Transactions	Transaction variables	Unsupervised anomaly detection + SHAP	Operational/benchmark; not firm-year splitting
Lei et al. (2023)	Transactions	Transaction variables	Distributed DNN	Holdout

(Source: The Researcher's Findings)

Despite growing interest in financial statement fraud detection, the literature still lacks a rigorous, apples-to-apples comparison of classical, deep, and hybrid models under a shared experimental protocol and leakage-resistant validation (e.g., the firm-level splitting and group-based cross-validation). Moreover, prior studies rarely isolated the incremental value of raw accounting items versus financial ratios or their combination making it difficult to disentangle algorithmic gains from feature-construction effects. Finally, evidence remains limited on whether a parsimonious feature subset can retain minority-class detection performance under extreme imbalance while improving practical deployability for audit planning and continuous monitoring.

Deep Learning and Machine Learning Methods for Fraud Detection

In recent years, the adoption of machine learning and advanced models particularly in the field of financial fraud detection has led to remarkable progress. In this study, a broad set of classical machine learning methods and advanced deep learning architectures were employed to analyze and accurately classify financial data. First, classical models including logistic regression, decision tree, random forest, and support vector machine were applied to establish baseline performance and evaluate traditional approaches.

In the deep learning component, a spiking classifier was first used, which due to its dynamic temporal nature can identify behavioral patterns that vary across financial periods (Jeyasothy et al., 2024). Next, residual networks ResNet-18 and ResNet-15 were employed. Then, a one-dimensional capsule network, as well as an efficient capsule network architecture, and a highly efficient and low-parameter variant of capsule networks were utilized for analyzing the financial data (Mazzia et al., 2021). In addition, the CAT-Net model was applied, integrating convolutional neural network layers, channel attention, and Transformers, enabling it to capture both local features of financial statements and long-term dependencies among different items (Islam et al., 2024). Finally, an advanced ensemble learning approach based on a Type-3 fuzzy decision-making system was used to intelligently and integratively combine the outputs of different models (Mehrabi Hashjin et al., 2024). Unlike simple aggregation methods, this system models the inherent uncertainty in financial data and fraud consequences, thereby enabling more accurate final decisions. To optimally tune parameters and fuzzy rules, an Improved Chaos Game Optimization algorithm which is an enhanced version of the Chaos Game Optimization algorithm was employed.

In detecting fraudulent financial transactions, Sai et al. (2023) evaluated LightGBM and XGBoost alongside deep neural networks. By applying measures such as hyperparameter tuning and imbalance-handling techniques, they showed that LightGBM achieved the best performance (accuracy ≈ 98.3%, F1 ≈ 0.70, AUC ≈ 0.96). Bao et al. (2020) developed a model for predicting fraud in U.S. publicly listed companies by using raw accounting numbers (rather than financial ratios) and applying ensemble learning. Their model significantly outperformed two widely used benchmarks: (1) the financial ratio–based logistic regression model proposed by Dechow et al. (2011), and (2) the SVM model based on ratios derived from raw data in Cecchini et al.'s study (2010).

Research Hypotheses

Based on the theoretical foundations presented, the research hypotheses were formulated as follows:

- **Hypothesis 1:** Deep learning models outperform classical machine learning baselines in detecting financial statement fraud.
- **Hypothesis 2:** Ensemble learning improves performance relative to single-model baselines under severe class imbalance.
- **Hypothesis 3:** A parsimonious feature subset can achieve performance comparable to or better than full feature sets, particularly in terms of F1-score.

Method

The present study was conducted to develop and evaluate methods for detecting fraud in financial statements. In terms of purpose, it is an applied study, and in terms of nature, it is descriptive. Moreover, because it uses historical data, it falls within the category of empirical and quasi-experimental research. The dependent variable is financial statement fraud. To enable comparability with prior studies, material misstatements reported in the U.S. SEC's AAERs were used as fraud cases, and the independent variables are financial ratios based on the dataset of Bao et al. (2020). Finally, by integrating the data, a new model with 24 features (23 financial ratios and one control variable) was developed for predicting and detecting fraud, allowing the performance of machine learning and deep learning methods to be evaluated in a manner comparable to previous research.

Research Variables

The dependent variable of the study is financial statement fraud, which is qualitative and nominal in nature and was coded as a binary variable (fraud firm = 1, non-fraud firm = 0). The fraud cases were identified based on material misstatements reported in the SEC's AAERs and in accordance with the dataset of Bao et al. (2020). The model inputs included 24 features, consisting of 23 financial ratios as independent variables and one control variable.

The independent variables were adopted from prior studies (e.g., Bao et al., 2020). After aligning the selection criteria with manifestations of fraudulent financial reporting, 23 financial ratios were selected and used, including: current ratio; current assets to total assets; fixed assets to total assets; working capital to total assets; total liabilities to total assets; retained earnings to total assets; sales to total assets; net income to total assets; gross profit to total assets; total liabilities to shareholders' equity; long-term debt to shareholders' equity; cost of goods sold to sales; gross profit to sales; net income to sales; operating expenses to sales; operating income to sales; net income to shareholders' equity; financial expenses to total liabilities; accounts receivable to sales; inventories to sales; cash and cash equivalents to total assets; inventories to total liabilities; and net income to gross profit.

Due to the role of financial distress in creating incentives and weakening the control environment, the Altman Z-score (1983) was included as a control variable in order to control for and examine the association between fraud and financial distress (Etemadi & Zolghi, 2013).

$$Z = 0.717x_1 + 0.847x_2 3.1x_3 + 0.42x_4 + 0.998x_5 \tag{1}$$

In Equation (1), x_1 denotes working capital to total assets, x_2 retained earnings to total assets, x_3 earnings before interest and taxes (EBIT) to total assets, x_4 book value of shareholders' equity to total liabilities, and x_5 sales to total assets.

In this study, advanced Shannon cross-entropy–based approach was employed for optimal feature selection in order to eliminate redundant information, improve model accuracy, and enhance computational efficiency. Ultimately, seven final features were selected. Moreover, using the dataset comprising 28 raw accounting variables and 14 financial ratios from Bao et al. (2020), the ratio-based results of this study were compared with the findings reported in that work. This dataset combines the raw data of Cecchini et al.'s study (2010) with financial ratios adapted from Dechow et al.'s study (2011) and Cecchini et al.'s study (2010). The full list of input variables is provided in Table 2.

Table 2.
The List of Input Variables

Input data used in Bao et al.'s study (2020, p. 213)			Input data used in the present study
The 28 raw financial variables adopted from Cecchini et al.'s study (2010):		Fourteen financial ratios (financial changes): including 11 ratios derived from Dechow et al.'s study (2011) and three ratios derived from Cecchini et al.'s study (2010).	The seven features used in this study were selected based on the theoretical literature and consisted of six financial ratios plus one control variable, and the Altman Z-score (1983).
- Short-Term Investments - Current Liabilities - Total Liabilities - Net Income - Property, Plant and Equipment - Preferred Stock - Retained Earnings - Accounts Receivable - Sales (Revenue) - Sale of Stock (Equity Issuance) - Taxes Payable - Total Taxes - Interest Expense - Inventory - Other Investments	- Current Liabilities - Total Assets - Common Shareholders' Equity - Cash and Short-Term Investments - Cost of Goods Sold - Common Shares Outstanding - Current Liabilities - Long-Term Debt Issuance - Total Long-Term Debt - Depreciation - Income Before Extraordinary Items	- Change in Common Stock - Change in Cash Margin - Change in Return on Assets - Securities Issuance - Book-to-Market Ratio - Discretionary Accruals Index - Adjusted Return on Equity - Soft Assets (excluding tangible assets) - Earnings Before Interest and Taxes Change in Free Cash Flow	- Current Ratio - Total Debt-to-Equity Ratio - Cost of Goods Sold to Sales Ratio - Operating Profit to Sales Ratio - Net Income to Shareholders' Equity Ratio - Financial Expenses to Total Debt Ratio - Financial Distress (Altman Z-score, 1983)

(Source: The Researcher's Findings)

Data Collection Method

The data used in this study were obtained from the Compustat accounting database via the dataset employed by Bao et al. (2020), which is publicly available on GitHub under the directory JarFraud/FraudDetection[1].

Statistical Sample

The statistical sample of this study included all firms listed on U.S. stock exchanges over the period 1991–2014, comprising a total of 146,045 observations. After removing missing data and applying preprocessing procedures, the final sample was reduced to 122,526 observations, including 121,624 observations in the non-fraud class and 902 observations in the fraud class.

Research Procedure

After collecting the financial ratios from the database, a preprocessing stage was conducted by removing missing values and preparing the data (including transforming them into images). The dataset was then split into training and testing sets. Next, machine learning–based and deep learning–based models were trained, and their hyperparameters were tuned to minimize error. Their performance was evaluated using the specified assessment metrics. Finally, the model with the highest accuracy was identified as the best-performing model for predicting and detecting the fraud.

Data Preprocessing Process

Before feeding the data into the machine learning models, a standardized preprocessing pipeline was applied. Specifically, observations containing missing values in any input variable were removed via listwise deletion. All predictors were then z-score standardized to place them on a common scale prior to model training and evaluation. Finally, we employed the mRMR feature selection method (Peng et al., 2005) to retain variables with the highest relevance to fraud identification while mitigating redundancy among predictors.

Converting Data into Images for Deep Learning

To enable the use of image-based deep architectures (e.g., ResNet-18, ResNet-50, and capsule networks), we transformed the tabular inputs into a two-dimensional time–frequency representation using the continuous wavelet transform (CWT) implemented in MATLAB. For each firm-year observation, the feature vector was treated as a short one-dimensional sequence. We computed the CWT under MATLAB's default parameterization, which employed the analytic Morse wavelet (symmetry parameter $\gamma = 3$, time-bandwidth product $P2 = 60$) with 10 voices per octave. The admissible scale range was selected automatically, and coefficients were L1-normalized. The transform yielded a complex coefficient matrix, from which we constructed the input image as the magnitude scalogram, $S = |cwt(x)|$. To ensure a uniform input resolution across observations and compatibility with standard CNN backbones, we resized the scalogram via interpolation (imresize) to a fixed image size (e.g., 224 × 224).

1. https://github.com/JarFraud/FraudDetection

Importantly, we did not replicate (tile) the original feature values, as repetition may introduce artificial periodicity and does not add information. Instead, resizing the scalogram preserved the original signal content while providing a consistent image representation for downstream training and evaluation.

Data Splitting

In this study, the data were first split into 80% for training and 20% for testing. To prevent information leakage and ensure a fair evaluation, the split was performed using a random, group-based partitioning strategy (based on the firm identifier), such that all observations for each firm were placed entirely in either the training set or the test set. Then, to reduce the risk of overfitting and obtain a more reliable estimate of the model performance, five-fold stratified, group-based cross-validation was applied within the training set; that is, while maintaining group independence in each fold, the class distribution (fraud/non-fraud) was kept as consistent as possible across folds.

Modeling Methods

To ensure that the empirical comparisons are conceptually motivated rather than a purely technical benchmark, we evaluated a set of model families that represent distinct learning mechanisms relevant to financial statement fraud detection. First, the classical tabular baselines logistic regression, the decision tree, random forest, and the support vector machine were included as established reference methods for structured accounting data. Second, deep feature-extraction models, a spiking neural network, and a one-dimensional convolutional neural network (1D-CNN) were considered to capture the non-linear patterns and local dependencies in the input series. Third, to examine whether the hierarchical representation learning benefits the fraud detection under an image-based representation, we evaluated CWT-derived image models, including ResNet-18, ResNet-50, a one-dimensional capsule network, and an efficient capsule network. Fourth, an attention-based Transformer architecture was included to model the global interactions across the input dimensions. Finally, we employed an ensemble learning strategy to aggregate complementary learners and improve robustness and performance stability under severe class imbalance.

Model Evaluation

The evaluation metrics used in this study included accuracy, precision, recall (sensitivity), specificity, and the F1-score.

Findings

Addressing the Class Imbalance Using Cost-Sensitive Learning

To address the severe class imbalance (the small number of fraud cases relative to non-fraud cases), a cost-sensitive learning approach, or class weighting, was applied throughout the entire modeling process. Accordingly, for the deep learning models, a weighted binary cross-entropy loss function was used:

$$L = -\left(w_{pos}\, y \log(p) + w_{neg}\,(1-y)\log(1-p)\right) \tag{2}$$

where $y \in \{0,1\}$ is the true label and p is the predicted probability for the fraud class (the positive class). The positive-class weight was determined based on the class imbalance ratio:

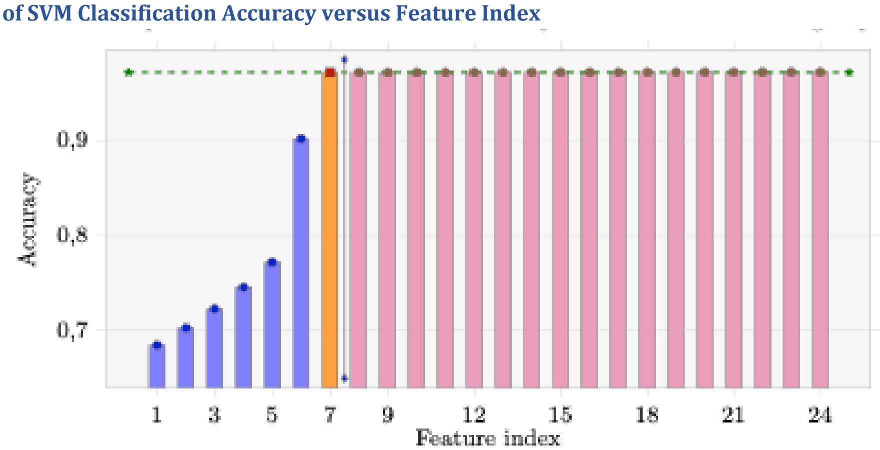
$$r = \frac{N_{neg}}{N_{pos}} \approx 134.8, \quad w_{pos} = \sqrt{r} \approx 11.6, \quad w_{neg} = 1 \tag{3}$$

Choosing $\sqrt{r}$ instead of r was intended to avoid overweighting the minority class and to maintain the stability of the learning process. For classical models, whose loss functions are not necessarily based on binary cross-entropy, the same idea was applied in an equivalent and implementable manner by incorporating class weights directly into the learning mechanism. Specifically, for logistic regression and other probabilistic models, a weighted cost function (i.e., class weighting) was used. For the support vector machine, the class weights were introduced into the error-penalty term so that misclassifications of the fraud class incurred a higher cost. For the decision tree and random forest models, the class weights were incorporated into split criteria and node impurity calculations (e.g., the Gini index or entropy), such that errors associated with the minority class exerted greater influence on split selection and the final model structure.

Feature Selection

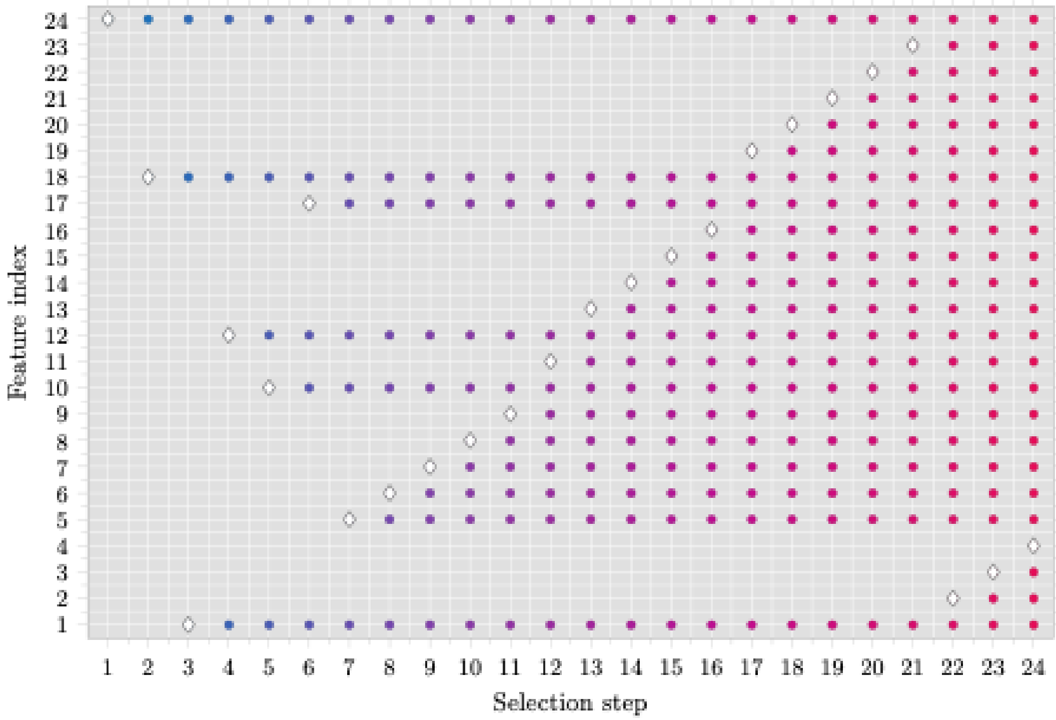
In this study, feature selection was performed using the mRMR method proposed by Peng et al. (2005). This method is based on Shannon mutual information and is designed to maximize the relevance of features to the target variable while minimizing redundancy among predictors. To determine the optimal number of input features, we conducted an incremental SVM-based evaluation, where features were added sequentially according to the mRMR ranking and the corresponding classification performance was tracked. Figure 1 reports the SVM classification accuracy as a function of the number of selected features, which provides an empirical basis for selecting a compact yet informative subset. Figure 2 visualizes the mRMR selection order across 24 steps and the cumulative selection path, illustrating how the candidate variables progressively enter the feature set. Based on the trade-off between predictive performance and parsimony, seven features were ultimately retained as the final input feature set.

Figure 1.
The Bar Chart of SVM Classification Accuracy versus Feature Index



(Source: The Researcher's Findings)

Figure 2.
The Feature Selection Order across 24 Steps

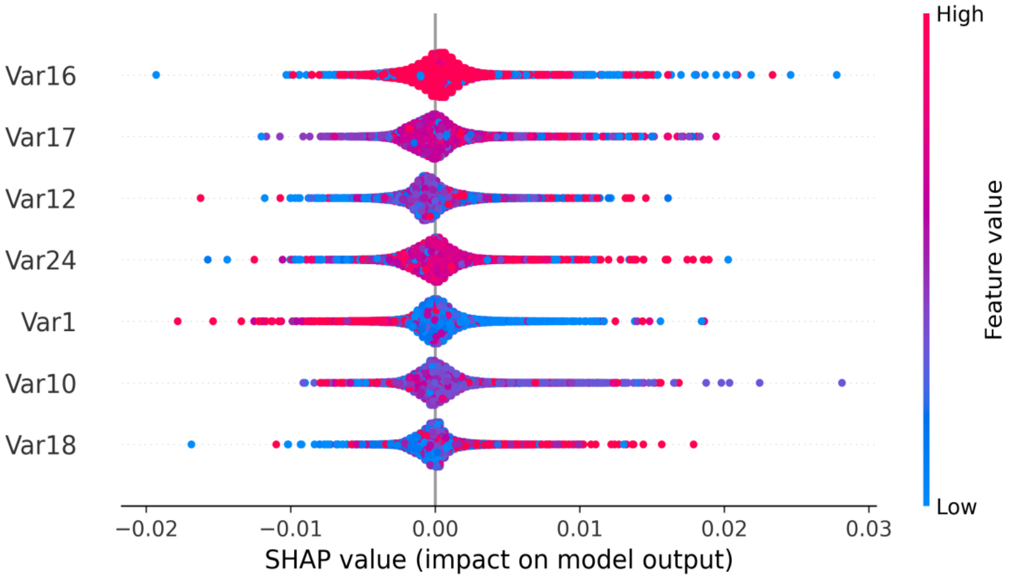


(Source: The Researcher's Findings)

The Model Interpretability and SHAP-based Explanation

To enhance the interpretability in a high-stakes setting such as financial statement fraud detection, we provided a model-based explanation of the Transformer’s decisions using SHAP analysis under the selected-feature setting. The SHAP summary (Figure 3) indicated that the model’s outputs are primarily driven by profitability-related measures most notably Operating Profit to Sales Ratio (Var16) and Net Income to Shareholders’ Equity (Return on Equity; Var17) which show the largest dispersion of SHAP values and therefore the greatest potential to shift predictions away from the baseline across observations. Financial health/distress, proxied by Altman’s Z-score (Var24), exhibited a comparatively clearer directional pattern, with higher values generally associated with positive SHAP contributions, whereas higher levels of the Current Ratio (Var1) more frequently align with negative contributions, suggesting that stronger short-term liquidity reduces the likelihood of the positive class in the learned decision function. The remaining indicators including Total Debt-to-Equity Ratio (Var10), Cost of Goods Sold to Sales Ratio (Var12), and Financial Expenses to Total Debt Ratio (Var18) displayed mixed contributions around zero, consistent with context-dependent and non-monotonic effects arising from non-linear interactions among profitability, leverage, liquidity, and distress signals captured by the attention-based architecture.

Figure 3.
The SHAP Beeswarm Summary for the Transformer



(Source: The Researcher's Findings)

Descriptive Statistics

Some descriptive statistics including mean, median, maximum, minimum, standard deviation, skewness, and kurtosis, are reported in Table 3. For example, the mean of the current ratio is 11.81. The standard deviation of this variable is 20.52, indicating that the average dispersion of the observations around the mean is of this magnitude. The median of the current ratio is 5.53, implying that 50% of the observations lie above and 50% lie below this value.

Table 3.
The Descriptive Statistics of the Study Dataset

Kurtosis	Skewness	Std.	Minimum	Maximum	Median	Mean	Feature
25.58	4.42	20.52	0.31	145.51	5.53	11.81	Current Ratio
36.12	5.44	5.35	0.05	41.43	1.07	2.45	Total Debt-to-Equity Ratio
58.59	7.33	2.38	0.07	21.06	0.67	1.00	Cost of Goods Sold to Sales Ratio
61.47	-7.55	1.43	-12.46	0.04	0	-0.23	Operating Profit to Sales Ratio
19.10	-0.76	1.29	-7.07	6.27	0.07	-0.03	Net Income to Shareholders' Equity
14.09	2.73	0.03	0	0.25	0.03	0.37	Financial Expenses to Total Debt Ratio
41.23	-5.83	10.58	-80.10	12.10	2.10	0.32	Financial Distress (Altman Z-score, 1983)

(Source: The Researcher's Findings)

Discussion

The dataset was first partitioned into an 80% training set and a 20% held-out test set using a firm-level group split to prevent the information leakage (i.e., all observations from a given firm were assigned exclusively to either training or test). The model development and hyperparameter tuning were conducted exclusively on the training set via five-fold stratified, firm-level group cross-validation. In each cross-validation iteration, the model was trained on four folds and validated on the remaining fold. After selecting the final configuration, the performance was assessed on the held-out test set. The results reported in Tables 4–7 summarize the test-set performance, aggregated over the five cross-validation–trained models.

According to the results reported in Tables 4–7, deep learning–based approaches clearly outperformed classical models. Among them, the Transformer neural network consistently achieved the best performance across all four datasets and provided the best balance between detecting fraud cases and controlling the false positive rate. After the Transformer, ensemble learning and the efficient capsule network ranked second and third, respectively (Tables 4–7). Among the convolutional models, the ResNet-50 showed the strongest performance within this family.

Finally, comparing the four datasets based on the best-performing model indicated that the highest F1-score was obtained with the 7 selected features (Table 7), followed by the combined 28+14 dataset (Table 6), the 28 raw variables (Table 4), and lastly the 14 financial ratios (Table 5). This pattern underscored that using raw data either alone or combined with financial ratios provides richer and complementary information for extracting the fraud-related patterns, leading to greater improvements in F1 compared with using financial ratios alone.

Practical Implications and Implementation in Auditing

From a practical standpoint, the proposed model should be viewed as a risk-screening tool that can be integrated into audit planning and continuous monitoring rather than as a stand-alone decision system. In an audit workflow, practitioners can compute the seven input features from routinely available financial statement data, generate a fraud-risk score using the trained model, and then use this score to (1) prioritize the firm-year engagements for enhanced procedures, (2) tailor the nature, timing, and extent of substantive testing in high-risk areas, and (3) support triage in regulatory surveillance. Implementation, however, requires explicit attention to operational constraints, including data quality and mapping consistency across reporting periods, threshold calibration to manage the cost of false positives, and periodic re-validation to mitigate the performance drift over time.

A further practical challenge is label and outcome latency in financial reporting fraud, ground-truth confirmation often arrives months or years after the reporting period (e.g., via enforcement actions), which complicates the timely model recalibration and performance monitoring.

Table 4.
The Performance of Twelve Methods on 28 Raw Input Variables Used in Bao et al. (2020)

Row	Method	Sensitivity	Precision	Specificity	Accuracy	F1-score
1	Decision Tree	0.28	0.06	0.96	0.96	0.09
2	Random Forest	0.34	0.08	0.97	0.96	0.13
3	Support Vector Machine	0.35	0.09	0.97	0.97	0.15
4	Spiking Neural Network	0.39	0.12	0.97	0.97	0.18
5	One-Dimensional Capsule Network	0.49	0.17	0.98	0.97	0.25
6	One-Dimensional Convolutional Neural Network	0.54	0.19	0.98	0.98	0.27
7	ResNet-18	0.52	0.18	0.98	0.97	0.26
8	ResNet-50	0.60	0.24	0.98	0.98	0.34
9	Efficient Capsule Network	0.67	0.29	0.98	0.98	0.40
10	Ensemble Learning	0.70	0.30	0.98	0.98	0.42
11	Transformer Neural Network	0.74	0.34	0.98	0.98	0.46
12	Logistic Regression	0.23	0.05	0.96	0.96	0.08

(Source: The Researcher's Findings)

Table 5.
The Performance of 12 Methods Used in This Study on 14 Financial Ratios from Bao et al. (2020)

Row	Method	Sensitivity	Precision	Specificity	Accuracy	F1-score
1	Decision Tree	0.25	0.05	0.96	0.95	0.08
2	Random Forest	0.32	0.07	0.97	0.96	0.12
3	Support Vector Machine	0.34	0.08	0.97	0.96	0.13
4	Spiking Neural Network	0.38	0.11	0.97	0.97	0.17
5	One-Dimensional Capsule Network	0.48	0.16	0.98	0.97	0.24
6	One-Dimensional Convolutional Neural Network	0.52	0.18	0.98	0.97	0.26
7	ResNet-18	0.49	0.17	0.98	0.97	0.25
8	ResNet-50	0.58	0.22	0.98	0.98	0.31
9	Efficient Capsule Network	0.65	0.27	0.98	0.98	0.38
10	Ensemble Learning	0.67	0.29	0.98	0.98	0.40
11	Transformer Neural Network	0.70	0.30	0.98	0.98	0.42
12	Logistic Regression	0.22	0.04	0.96	0.95	0.07

(Source: The Researcher's Findings)

Table 6.
The Performance of 12 Methods Used in This Study on Combined Set of 28 Raw Variables and 14 Financial Ratios from Bao et al. (2020)

Row	Method	Sensitivity	Precision	Specificity	Accuracy	F1-score
1	Decision Tree	0.29	0.06	0.96	0.96	0.10
2	Random Forest	0.35	0.09	0.97	0.96	0.15
3	Support Vector Machine	0.38	0.10	0.97	0.97	0.16
4	Spiking Neural Network	0.41	0.13	0.97	0.97	0.20
5	One-Dimensional Capsule Network	0.52	0.19	0.98	0.98	0.27
6	One-Dimensional Convolutional Neural Network	0.55	0.21	0.98	0.98	0.30
7	ResNet-18	0.54	0.2	0.98	0.98	0.29
8	ResNet-50	0.61	0.25	0.98	0.98	0.35
9	Efficient Capsule Network	0.70	0.30	0.98	0.98	0.42
10	Ensemble Learning	0.71	0.32	0.98	0.98	0.45
11	Transformer Neural Network	0.76	0.36	0.98	0.98	0.48
12	Logistic Regression	0.24862	0.05	0.96	0.96	0.09016

(Source: The Researcher's Findings)

Table 7.
The Performance of 12 Methods Used in This Study on Seven Features

Row	Method	Sensitivity	Precision	Specificity	Accuracy	F1-score
1	Decision Tree	0.32	0.06	0.96	0.96	0.11
2	Random Forest	0.38	0.1	0.97	0.97	0.15
3	Support Vector Machine	0.39	0.10	0.97	0.97	0.17
4	Spiking Neural Network	0.44	0.13	0.97	0.97	0.21
5	One-Dimensional Capsule Network	0.55	0.2	0.98	0.98	0.29
6	One-Dimensional Convolutional Neural Network	0.60	0.22	0.98	0.98	0.32
7	ResNet-18	0.58	0.21	0.98	0.98	0.31
8	ResNet-50	0.65	0.27	0.98	0.98	0.38
9	Efficient Capsule Network	0.71	0.32	0.98	0.98	0.44
10	Ensemble Learning	0.75	0.34	0.98	0.98	0.46
11	Transformer Neural Network	0.77	0.38	0.99	0.98	0.51
12	Logistic Regression	0.28	0.06	0.96	0.96	0.09

(Source: The Researcher's Findings)

These findings were derived from U.S. listed firms with AAER-identified misstatements; therefore, the implications should be interpreted within the U.S. regulatory and reporting environment.

Table 8.
The McNemar Paired Comparisons versus the Transformer

Model	Difference (%)	95% CI for Difference (%)	p_value
Decision Tree	-2.54	[-2.634, -2.456]	0
Random Forest	-1.88	[-1.965, -1.811]	0
Support Vector Machine	-1.72	[-1.794, -1.647]	0
Spiking Neural Network	-1.33	[-1.403, -1.271]	0
One-Dimensional Capsule Network	-0.87	[-0.927, -0.818]	4.3757e-263
One-Dimensional Convolutional Neural Network	-0.69	[-0.742, -0.643]	1.452e-197
ResNet-18	-0.76	[-0.819, -0.716]	9.5197e-225
ResNet-50	-0.45	[-0.497, -0.414]	2.9818e-116
Efficient Capsule Network	-0.23	[-0.268, -0.195]	4.2567e-37
Ensemble Learning	-0.15	[-0.197, -0.120]	3.2608e-16
Logistic Regression	-2.68	[-2.779, -2.597]	0

(Source: The Researcher's Findings)

Table 8 reports paired, two-sided exact McNemar tests comparing each model with the Transformer under the seven-feature setting. Difference (%) represents the net paired difference in instance-level correctness derived from the discordant pairs. Negative values indicate that the Transformer correctly classifies more cases than the comparator model. Across all baselines, the estimated differences were negative and the 95% confidence intervals lay entirely below zero, and all exact tests were significant ($p<0.001$), indicating a consistent instance-level advantage for the Transformer. The magnitude of the advantage was largest for classical tabular baselines and narrowed for stronger deep/hybrid models, yet remained statistically significant throughout.

Practical Integration into Audit Planning and Regulatory Surveillance

The proposed Transformer model can be integrated as a risk-scoring layer in audit planning to rank firms by predicted fraud likelihood, thereby improving the resource allocation and the design of targeted substantive procedures (e.g., focusing on revenue-related accounts and accrual-intensive items). The parsimonious feature subset enables fast, low-cost screening and allows the score to be mapped into the operational risk tiers (low/medium/high) that trigger predefined audit responses and documentation requirements. For regulators, the same score supports continuous surveillance, early-warning alerts, and prioritization of issuers for review under constrained supervisory capacity. Coupling the risk score with explainability outputs (e.g., feature-attribution summaries) further enhances the interpretability and practitioner trust in the resulting flags.

Conclusion

This study developed and evaluated a data-driven framework for detecting financial statement fraud using a unified, leakage-resistant evaluation protocol and a broad benchmark set of classical, deep, and hybrid classifiers. Using AAER-identified fraud cases matched to Compustat accounting data for U.S. listed firms (1991–2014), we compared four input scenarios including raw accounting items, financial ratios, their combination, and a compact seven-feature subset (six ratios plus Altman's Z-score). Across all scenarios, deep and hybrid models consistently outperformed the classical baselines, underscoring the non-linear and interaction-driven structure of fraud signals in financial reporting data. The Transformer model delivered the strongest and most stable performance, achieving the best overall balance between fraud detection capability and false-positive control, with its highest F1-score obtained under the parsimonious seven-feature configuration. Importantly, the combined raw-item and ratio representation outperformed the ratios-only configuration, indicating that raw financial statement items provide incremental discriminatory information beyond the engineered ratios.

Beyond the predictive accuracy, the study emphasized deployability in audit and regulatory settings. The selected seven-feature subset enabled low-cost screening based on routinely available financial statement information, while the SHAP-based interpretability analysis provided transparency on which accounting signals most strongly influence the Transformer's decisions. Together, these results supported the use of the proposed model as a risk-screening layer to prioritize firm-year engagements for further investigations and to inform audit planning and supervisory surveillance.

Limitations and Future Research Directions

The limitations of this paper include missing or inconsistent data as well as class imbalance between the fraudulent and non-fraudulent classes. In addition, a systematic comparison between the imbalance-handling strategy adopted in this study, cost-sensitive learning and alternative data-imbalance techniques such as the Synthetic Minority Over-sampling Technique (SMOTE) and Adaptive Synthetic Sampling (ADASYN), as well as potential hybrid schemes combining SMOTE/ADASYN with cost-sensitive learning, would be a valuable extension for future research in financial statement fraud detection.

REFERENCES

Achakzai, M. A. K., & Juan, P. (2022). Using machine learning Meta-Classifiers to detect financial frauds. *Finance Research Letters*, *48*, 102915. https://doi.org/10.1016/J.FRL.2022.102915.

Ahmed, K., & Courtis, J. K. (2015). The determinants of financial ratio disclosures and quality: Evidence from an emerging market. *British Accounting Review*, *31*(1), 35–61. https://doi.org/10.1006/BARE.1998.0082.

Ahmed, K. H., Axelsson, S., Li, Y., & Sagheer, A. M. (2025). A credit card fraud detection approach based on ensemble machine learning classifier with hybrid data sampling. *Machine Learning with Applications*, *20*, 100675. https://doi.org/10.1016/J.MLWA.2025.100675.

Altman, E.I. (1983) Corporate Financial Distress: A Complete Guide to Predicting, Avoiding, and Dealing with Bankruptcy. Wiley, New York (1983), 368.

Arboleda, F. J. M., Guzman-Luna, J. A., & Torres, I. D. (2018). Fraud detection-oriented operators in a data warehouse based on forensic accounting techniques. *Computer Fraud & Security*, *2018*(10), 13–19. https://doi.org/10.1016/S1361-3723(18)30098-8.

Azim Mim, M., Majadi, N., & Mazumder, P. (2024). A soft voting ensemble learning approach for credit card fraud detection. *Heliyon*, *10*(3), e25466. https://doi.org/10.1016/j.heliyon.2024.e25466.

Bao, Y., Ke, B., Li, B., Yu, Y. J., & Zhang, J. (2020). Detecting accounting fraud in publicly traded U.S. firms using a machine learning approach. *Journal of Accounting Research*, *58*(1), 199–235. https://doi.org/10.1111/1475-679X.12292.

Bhattacharya, I., & Mickovic, A. (2024). Accounting fraud detection using contextual language learning. *International Journal of Accounting Information Systems*, *53*, 100682. https://doi.org/10.1016/J.ACCINF.2024.100682.

Cai, S., & Xie, Z. (2024). Explainable fraud detection of financial statement data driven by two-layer knowledge graph. *Expert Systems with Applications*, *246*, 123126. https://doi.org/10.1016/J.ESWA.2023.123126.

Cao, R., Wang, J., Mao, M., Liu, G., & Jiang, C. (2023). Feature-wise attention based boosting ensemble method for fraud detection. *Engineering Applications of Artificial Intelligence*, *126*, 106975. https://doi.org/10.1016/J.ENGAPPAI.2023.106975.

Cecchini, M., Aytug, H., Koehler, G. J., & Pathak, P. (2010). Detecting management fraud in public companies. *Management Science*, *56*(7), 1146–1160. https://doi.org/10.1287/MNSC.1100.1174.

Dechow, P. M., Ge, W., Larson, C. R., & Sloan, R. G. (2011). Predicting material accounting misstatements. *Contemporary Accounting Research*, *28*(1), 17–82. https://doi.org/10.1111/J.1911-3846.2010.01041.X.

Etemadi, H., & Zolghi, H. (2013). Application of logistic regression in detecting fraudulent financial reporting. *Danesh-e Hesabresi (Auditing Knowledge)*, *13*(51) 145-163.

Islam, M. R., Qaraqe, M., Qaraqe, K., & Serpedin, E. (2024). CAT-Net: Convolution, attention, and transformer based network for single-lead ECG arrhythmia classification. *Biomedical Signal Processing and Control*, *93*, 106211. https://doi.org/10.1016/J.BSPC.2024.106211.

Jeyasothy, A., Suresh, S., Ramasamy, S., & Sundararajan, N. (2024). Development of a novel transformation of spiking neural classifier to an interpretable classifier. *IEEE Transactions on Cybernetics*, *54*(1), 3–12. https://doi.org/10.1109/TCYB.2022.3181181.

Kanapickienė, R., & Grundienė, Ž. (2015). The model of fraud detection in financial statements by means of financial ratios. *Procedia - Social and Behavioral Sciences*, *213*, 321–327.

https://doi.org/10.1016/J.SBSPRO.2015.11.545.

Karnavou, E., Cascavilla, G., Marcelino, G., & Geradts, Z. (2025). I know you're a fraud: Uncovering illicit activity in a Greek bank transactions with unsupervised learning. *Expert Systems with Applications*, *288*, 128148. https://doi.org/10.1016/J.ESWA.2025.128148.

Kim, Y. J., Baik, B., & Cho, S. (2016). Detecting financial misstatements with fraud intention using multi-class cost-sensitive learning. *Expert Systems with Applications*, *62*, 32–43. https://doi.org/10.1016/J.ESWA.2016.06.016.

Lei, Y. T., Ma, C. Q., Ren, Y. S., Chen, X. Q., Narayan, S., & Huynh, A. N. Q. (2023). A distributed deep neural network model for credit card fraud detection. *Finance Research Letters*, *58*, 104547. https://doi.org/10.1016/J.FRL.2023.104547.

Lu, J., Xu, Q., & Hu, J. (2026). A novel graph learning framework for interpretable and imbalance financial fraud detection. *Engineering Applications of Artificial Intelligence*, *167*, 113709. https://doi.org/10.1016/J.ENGAPPAI.2025.113709.

Mazzia, V., Salvetti, F., & Chiaberge, M. (2021). Efficient-CapsNet: capsule network with self-attention routing. *Scientific Reports, 11*(1), 14634. https://doi.org/10.1038/s41598-021-93977-0.

Mehrabi Hashjin, N., Amiri, M. H., Mohammadzadeh, A., Mirjalili, S., & Khodadadi, N. (2024). Novel hybrid classifier based on fuzzy type-III decision maker and ensemble deep learning model and improved chaos game optimization. *Cluster Computing*, *27*(7), 10197–10234. https://doi.org/10.1007/S10586-024-04475-7/METRICS.

Narayana Gorle, V. L., & Panigrahi, S. (2026). An efficient heuristic optimization-based fraudulent activities detection in the financial sector using adaptive machine learning and deep learning system. *Expert Systems with Applications*, *302*, 130551. https://doi.org/10.1016/J.ESWA.2025.130551.

Peng, H., Long, F., & Ding, C. (2005). Feature selection based on mutual information: Criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *27*(8), 1226–1238. https://doi.org/10.1109/TPAMI.2005.159.

Sai, C. V., Das, D., Elmitwally, N., Elezaj, O., & Islam, M. B. (2023). *Explainable ai-driven financial transaction fraud detection using machine learning and deep neural networks*. https://doi.org/10.2139/SSRN.4439980.

Shao, Z., Yu, H., Wen, J., Liu, Z., & Qi, P. (2026). A graph fraud detection model based on mutual information. *Neurocomputing*, *663*, 131972. https://doi.org/10.1016/J.NEUCOM.2025.131972.

Vakilifard, H. R., Jabarzadeh Kangarlouei, S., & Pourreza Sultan Ahmadi, A. (2009). An investigation of the characteristics of fraud in financial statements. *Monthly Magazine of the Iranian Association of Certified Public Accountants*, (210), 26–41.

Zhang, Z., Wang, Z., & Cai, L. (2025). Predicting financial fraud in Chinese listed companies: An enterprise portrait and machine learning approach. *Pacific-Basin Finance Journal*, *90*, 102665. https://doi.org/10.1016/J.PACFIN.2025.102665.

Journal of
Knowledge Economy Studies

https://kes.hmu.ac.ir

Hazrat-e Masoumeh University

Intelligent Risk Processing and Opportunity Formation in Financial Markets: The Superior Performance of the HERC Algorithm in Efficient Portfolio Construction

Article Type:
Research Article

Mahsa Safavi Iranji[1]
Faculty of Social Sciences and Economics, Alzahra University, Tehran, Iran
E-mail: mahsa.safavi@ut.ac.ir

Majid Zanjirdar[2*]
Faculty of Social Sciences and Economics, Alzahra University, Tehran, Iran
Corresponding Author
E-mail: majid.zanjirdar@iau.ir

Received: 15 November 2025
Revised: 25 December 2025
Accepted: 27 January 2026
Available Online: 28 January 2026

Abstract

Hierarchical portfolio optimization methods, particularly the Hierarchical Equal Risk Contribution (HERC) approach, have become increasingly prominent in financial research due to their effectiveness in balancing risk and enhancing diversification. Unlike traditional methods such as Equal-Weight (EW) and Inverse Volatility (IV), which rely on oversimplified assumptions and often underperform in volatile markets, HERC allocates capital by distributing risk more efficiently across assets. This study examines the performance of the HERC model relative to EW and IV to determine its ability to convert risk into investment opportunities under fluctuating market conditions. The methodology follows a structured process that includes deriving variables from multiple data sources, conducting thorough data cleaning and normalization, and implementing traditional allocation models as benchmarks. Advanced hierarchical clustering techniques are then applied to provide a more innovative allocation framework. Rigorous hypothesis testing is used to validate the results, and portfolio performance is evaluated using established statistical metrics. Findings reveal that HERC—especially its single linkage and average linkage versions—delivers substantially higher risk-adjusted returns, as measured by the Sharpe and Sortino ratios, compared to EW and IV. The proposed methodology not only improves overall investment outcomes but also enables more effective risk and return management, making it a strong alternative to conventional portfolio construction and risk evaluation approaches.

Keywords

Asset allocation, Hierarchical equal risk contribution (HERC), Risk management, Unsupervised learning.

Cite this article: Safavi Iranji, M. & Zanjirdar, M., (2026). Intelligent risk processing and opportunity formation in financial markets: the superior performance of the HERC algorithm in efficient portfolio construction. *Journal of Knowledge Economy Studies (JKES)*, 3(1), 25-46. DOI: http//doi.org/10.22034/kes.2026.2077852.1088

Online ISSN: 3060-7329

Introduction

In the volatile financial world, risk management has always been a significant challenge. Risk is traditionally viewed as a threat to returns; however, the ability to transform risk into opportunity is key to successful investing. The increasing integration of global financial markets necessitates advanced quantitative models (Rostami, et al 2025). In recent years, advanced algorithms have increasingly played a crucial role in enhancing investment management processes. Despite notable advancements in asset allocation algorithms, many traditional methods, such as the Markowitz model, still lack adequate efficiency in addressing real-world complexities due to their simplifying assumptions (Markowitz, 1952). Conversely, innovative approaches like the Hierarchical Equal Risk Contribution (HERC) algorithm have recently emerged, but comprehensive research comparing its performance with other advanced asset allocation models remains limited (Raffinot, 2018). Although significant theoretical and practical progress has been made in asset allocation, traditional models continue to face challenges such as complex asset correlations and imbalanced risk distribution in portfolios (Huang & Gao, 2021). These limitations not only lead to inefficient portfolios but also hinder the potential to improve risk-adjusted returns. The advent of new technologies, including machine learning in asset management, has created unparalleled opportunities for refining asset allocation methods (Schwendner et al., 2021; Menvouta et al., 2023). This research endeavor seeks to assess the efficacy of the Hierarchical Equal Risk Contribution (HERC) algorithm as an innovative method for risk management and asset distribution. The research seeks to assess HERC's performance compared to traditional and innovative models, identify its advantages and limitations, and provide recommendations for enhancing its efficiency in real market conditions. The HERC algorithm, introduced by Thomas Raffinot, represents a cutting-edge approach to portfolio risk management. By leveraging hierarchical clustering and emphasizing balanced risk distribution, HERC effectively manages risk while delivering more stable returns (Raffinot, 2018). As financial markets grow increasingly complex and economic crises emerge, investors seek tools that not only ensure sustainable returns but also guarantee balanced risk distribution (Millea & Edalat, 2022). The structure of this paper is organized as follows: the first section examines the theoretical foundations of the study, introducing the concept of hierarchical equal risk contribution portfolios and analyzing methodologies proposed by Lopez de Prado and Raffinot (Raffinot, 2017; López de Prado, 2016). Subsequent sections review the research background, discuss the dataset and associated challenges, and explain the research methodology. Ultimately, the findings derived from the empirical analyses will be presented and interpreted. This research endeavors to validate the HERC algorithm's ability to convert risk into opportunity and construct efficient portfolios. The findings are expected to open new avenues for investors and researchers in risk management and asset allocation while addressing existing gaps in the literature on asset allocation (Sen & Mehtab, 2021; Sajadi et al., 2024). The necessity of this study lies in addressing real market challenges and leveraging recent advancements in financial technology.

Theoretical Framework

Risk management is a process in which investors and financial managers identify, evaluate, and manage risks associated with assets and investments (Nourahmadi & Sadeqi, 2021). This concept has been thoroughly explored in both classical and modern financial theories, such as Markowitz's portfolio theory (Markowitz, 1952). Asset allocation involves distributing capital across various asset classes (e.g., stocks, bonds, and alternative assets) to optimize returns while managing risk. Within this domain, traditional models like Minimum Variance and advanced models like Hierarchical Equal Risk Contribution (HERC) hold particular significance (Raffinot, 2018). The HERC algorithm is a novel approach to portfolio management that leverages hierarchical clustering techniques to balance risk distribution among assets. By utilizing hierarchical structures, HERC addresses the limitations of traditional models such as Markowitz's and equitably distributes risks across the portfolio. Modern Portfolio Theory (MPT), introduced by Harry Markowitz, is founded on the risk-return hypothesis, proposing that portfolio optimization should maximize expected returns while minimizing risk. Additionally, models like Equal Risk Contribution (ERC) and parametric models aim to distribute risks evenly among assets (Maillard et al., 2010). However, these models often struggle with the complexities of asset correlations, paving the way for clustering-based methods like HERC. Hierarchical clustering methods, classified as unsupervised machine learning techniques, utilize asset correlation data to create tree-like structures for grouping assets (Duarte & De Castro, 2020). These structures, employed in the HERC algorithm, enable the identification of complex relationships among assets, thereby improving asset allocation efficiency and portfolio optimization. While traditional theories such as Markowitz's model and risk-balancing approaches like ERC hold a prominent place in financial literature, their limitations—such as the assumption of linear correlations and the inability to handle complex market structures—highlight the importance of algorithms like HERC. Despite the growing relevance of HERC in asset allocation, comprehensive studies analyzing its performance and comparing it with other models remain scarce, creating a critical gap in financial literature (Huang & Gao, 2021).

1. INVERSE VARIANCE PORTFOLIO

The logic behind the Inverse Variance (IV) portfolio allocation strategy is fundamentally straightforward: in the IV strategy, risk is measured by variance, and assets are weighted inversely proportional to their variances. Consequently, based on the standard deviation of returns, the IVP assigns the following weights to N assets (Ferretti, 2022):

$$w_{IV} = \frac{\frac{1}{\sigma_i^2}}{\sum_{i=1}^{N} \frac{1}{\sigma_i^2}} \tag{1}$$

where:

wi: Weight of asset i in the portfolio.

σ_i^2: Variance of the returns of asset i.

N: Total number of assets in the portfolio.

This method ensures that assets with lower variance receive higher weights, aligning the portfolio's construction with the principle of minimizing risk by favoring less volatile assets.

The primary advantage of this method lies in its computational simplicity. However, its accuracy and efficiency may diminish when faced with limitations and assumptions, such as the independence of assets or the exclusion of expected returns. Consequently, in practice, more advanced methods and models may be required to construct an optimal investment portfolio.

2. EQUAL-WEIGHT PORTFOLIO

In this straightforward method, each asset is assigned an equal weight. This approach is effective when the returns of all assets are entirely uncertain and random, with no subjective biases or preferences influencing the allocation.

$$w_i = \frac{1}{n} \tag{2}$$

The portfolio weight vector is defined as:

$$w_{EW} = \left(\frac{1}{n}, \dots, \frac{1}{n}\right)^T \tag{3}$$

DeMiguel et al. (2009) made noteworthy findings, showing that despite its simplicity, the equal-weight portfolio often outperforms more complex strategies (DeMiguel et al., 2009). Specifically, the authors found that among 14 evaluated models, equal-weight portfolios consistently demonstrated superior performance in terms of Sharpe ratio, actual returns, and turnover in out-of-sample backtests. They argue that the benefits of diversification outweigh estimation errors, causing many sophisticated methods to underperform relative to this simple allocation strategy consistently.

3. HIERARCHICAL RISK PARITY

Hierarchical Risk Parity (HRP) is a portfolio optimization approach that is applied through a three-step process. The hierarchical clustering method is implemented using the Scipy library. Below is a description of each step in terms of its implementation stages.

Hierarchical Clustering

First, the correlation matrix is used to form clusters using the agglomerative hierarchical clustering algorithm. This ensures that assets within clusters are as similar as possible in terms of risk characteristics.

Sorting the Matrix

Second, the assets are sorted on the covariance matrix in such a way that similar assets are grouped closer together, facilitating a more balanced risk distribution across the portfolio.

Recursive Bisection

Third, in the recursive bisection stage, the sorted covariance matrix is used to allocate weights in such a way that the clusters achieve equal sizes. This process involves repeatedly halving the covariance matrix into sub-clusters until each asset is uniquely assigned to a cluster. To initiate this process, the algorithm starts with the following initialization.

Initialization of Clusters and Asset Weights:

Define the list of clusters as C0 = {C0} where C0 = {C0}n = 1,...,N i.e., all assets start in a single cluster.

Set the initial asset weights as $w_n = 1$, for all n∈[1,...,N]

Termination Condition:

If $|C_i| = 1$ for all Ci ∈ C , stop.

Cluster Division:

For each cluster Ci ∈ C, where, $|C_i| > 1$, continue.

Recursive Bisection:

Divide each cluster Ci into two subclusters Ci1 ∪ Ci2 , such that $|C_i^1| = \mathrm{int}\left[\frac{1}{2}|C_i|\right]$.

Define the weights within each subcluster Cij as:

$$w_i^{(j)} = \frac{tr\left[\hat{\Sigma}_i^{(j)}\right]^{-1}}{\sum_i tr\left[\hat{\Sigma}_i^{(j)}\right]-1} \quad for \quad j = 1,2 \tag{4}$$

where $\hat{\Sigma}_i^{(j)}$ is the covariance matrix for the subcluster C_{ij}.

Define and calculate the variance for each subcluster C_{ij} as:

$$V_i^j = w_i^{(j)\mathrm{T}} \hat{\Sigma}_i^{(j)} w_i^{(j)} \quad for \quad j = 1,2 \tag{5}$$

Compute the partition factor α_1 and α_2 as:

$$\alpha_1 = 1 - \frac{V_i^1}{V_i^1 + V_i^2} \ , \quad \alpha_2 = 1 - \alpha_1 \tag{6}$$

Rescale the asset allocations w_n for all n ∈ Ci1 by factor α1.

Rescale the asset allocations w_n for all n ∈ Ci2by factor α2.

Repeat step 2 in a loop until the termination condition is met.

This description outlines the key steps in the Hierarchical Risk Parity (HRP) algorithm, which employs hierarchical clustering to balance risk across different assets and groups them recursively until each cluster contains a single asset. The process ensures that assets with higher correlation are grouped together and that risk is allocated in a balanced manner within each cluster (Deković & Šimović, 2025; Jain & Jain, 2019).

4. HIERARCHICAL EQUAL RISK CONTRIBUTION

The HERC method is a portfolio optimization technique based on hierarchical clustering and equal risk contribution across clusters of assets. It consists of a four-step process, which is similar to the Hierarchical Risk Parity (HRP) method but includes an additional step to determine the optimal number of clusters. It also differs in the assumptions regarding recursive halving. The four stages of the HERC method are as follows:

(1) Hierarchical Correlation Clustering

In this first step, a correlation matrix is transformed into a distance matrix. The distance matrix is then used to form clusters via a hierarchical agglomerative clustering algorithm. This process follows these steps:

a) Convert the Correlation Matrix:

The correlation matrix, which captures the relationship between asset returns, is converted into a distance matrix. The distance between two assets is generally computed as:

$$\text{Distance } (i, j) = 1 - corr(i, j) \tag{7}$$

b) Hierarchical Clustering:

Using the distance matrix, hierarchical clustering is applied. This technique starts with each asset as its own cluster and iteratively merges clusters based on their similarity (minimized distance) until all assets are in a single cluster.

c) Hierarchical Structure:

The resulting hierarchical structure or dendrogram provides a visual representation of asset relationships, which is crucial for the next steps.

(2) Determining the Optimal Number of Clusters

Determining the optimal number of clusters after hierarchical clustering is crucial, as it shapes the portfolio's structure by specifying sub-clusters. Methods like Silhouette Score, Gap Statistic, and Elbow Method are used to identify the ideal cluster count. This choice balances risk distribution depth with diversification, ensuring an effective allocation strategy.

(3) Recursive Hierarchical Halving

In this step, the sorted covariance matrix (from the clustering result) is recursively halved. The goal is to split the matrix into sub-clusters based on the hierarchical structure until the optimal number of clusters is achieved. This process includes:

a) Recursive Halving:

Starting with the covariance matrix, the assets are recursively divided into sub-clusters based on the hierarchical tree. This continues until each cluster has been divided into an optimal number of sub-clusters

b) Assigning Weights:

For each resulting sub-cluster, the weights are assigned to assets by balancing the risk contribution across all assets within that sub-cluster. The weights are calculated iteratively by considering the risk (variance) of each asset and its contribution to the overall cluster risk. The final weights are adjusted by the risk contribution factor, which ensures that the risk is equally distributed among all assets.

To implement the Hierarchical Equal Risk Contribution (HERC) algorithm, the process begins with initializing the algorithm and iteratively refining the asset weights through a recursive halving and equal risk distribution procedure. Below is the step-by-step implementation.

Step (1): Initialization

1. Initialize Clusters:

Start by setting the list of items (assets) as C0 = {C0} with C0 = {n}, n = 1,…,N. Here, N is the total number of assets in the portfolio.

2. Set Initial Weights:

The initial weights w_n for each asset are set to 1:

$$w_n = 1, \forall n \in [1, \cdots, N] \tag{8}$$

This assumes an equal initial distribution of weights across all assets.

Step (2): Recursive Clustering and Halving Process

For each iteration i = 0 to k* - 1, where k* is a predetermined number of clusters:

3. Ensure the Cluster Has More Than One Asset:

The algorithm continues only for clusters Ci where Ci > 1, i.e., the cluster has more than one asset.

4. Split the Clusters:

Divide each cluster C_i into two sub-clusters: Ci1 (left sub-cluster) and Ci2 (right sub-cluster).

5. Assign Initial Weights for the Sub-clusters:

The weights inside each sub-cluster Ci(j) for j= 1,2 are assigned as follows:

$$w_i^{(j)} = \frac{tr\left[\hat{\Sigma}_i^{(j)}\right]^{-1}}{\sum_i tr\left[\hat{\Sigma}_i^{(j)}\right]-1} \tag{9}$$

Here, $\hat{\Sigma}_i^{(j)}$ is the covariance matrix of sub-cluster Cij.

6. Calculate Variance for Each Sub-cluster:

The variance of each sub-cluster $V_i^{(j)}$ for j = 1.2 is computed as:

$$v_i^j = w_i^{(j)\mathrm{T}} \hat{\Sigma}_i^{(j)} w_i^{(j)} \tag{10}$$

7. Calculate the Split Factor:

The factor α1 for the left sub-cluster and α2 for the right sub-cluster are calculated using the risk contributions of each sub-cluster:

$$\alpha_1 = 1 - \frac{RC_i^1}{RC_i^1 + RC_i^2}, \quad \alpha_2 = 1 - \alpha_1 \tag{11}$$

where RCi1 and RCi2 are the risk contributions of the left and right sub-clusters, respectively.

8. Rescale Weights for Each Sub-cluster:

Rescale the weights inside each sub-cluster:

$$\begin{aligned} &w_n \quad \forall n \in C_i^1 \quad by \quad \alpha_1 \\ &w_n \quad \forall n \in C_i^2 \quad by \quad \alpha_2 \end{aligned} \tag{12}$$

Step (3): Cluster Weighting

Once the clusters have been divided and the weights adjusted, the final weights between and within clusters are computed.

1. Within-Cluster Weights:

The weights within each cluster are calculated using an inverse variance allocation:

$$w_i^* = \frac{tr\left[\hat{\Sigma}_i\right]^{-1}}{\sum_i tr\left[\hat{\Sigma}_i\right]-1} \tag{13}$$

where $\hat{\Sigma}_i$ represents the covariance matrix for each cluster C_i.

2. Final Asset Weights:

The final asset weights are obtained by multiplying the weights between clusters with the weights within clusters:

$$w_{HERC} = W.w \tag{14}$$

where:

W is the vector of weights between clusters, and w is the vector of weights within each cluster. The result ensures that the total portfolio weight sums to 1 (Raffinot, 2018; Deković & Šimović, 2025; Sjöstrand et al., 2020).

5. TYPES OF LINKAGE METHODS IN HIERARCHICAL CLUSTERING

A crucial consideration in hierarchical clustering is the selection of the distance metric and linkage method. The distance metric determines how the distance between two data points is measured, while the linkage method defines how the distances between clusters are combined to calculate the overall distance between them. Commonly used distance metrics include Euclidean distance, Manhattan distance, and cosine similarity, while common linkage methods include complete linkage, single linkage, and average linkage (Zhao & Karypis, 2002).

Hierarchical clustering is one of the widely used methods in data analysis aimed at identifying natural groupings (clusters) within data. This method reveals the underlying structure of data by creating a hierarchy of clusters. Among the various hierarchical clustering methods, Single Link, Average Link, and Ward's method are particularly popular due to their simplicity and effectiveness. In this research, we focus on the Single Link, Average Link, and Ward's methods. These methods form clusters using different distance metrics, each with its own advantages and disadvantages.

1. Single Link Method

In the Single Link method, the distance between two clusters is defined as the minimum distance between the nearest pair of points from the two clusters. In other words, at each step, the two clusters with the closest points are merged. This method is less sensitive to the shape of the clusters and can identify clusters with irregular shapes. However, it is prone to forming chain-like clusters, which may lead to counterintuitive results.

2. Average Link Method

In the Average Link method, the distance between two clusters is defined as the average distance between all pairs of points from the two clusters. This method is less prone to creating chain-like clusters compared to the Single Link method and tends to produce more spherical clusters. However, its sensitivity to the size of the clusters is a limitation.

3. Ward's Method

Ward's method aims to minimize the within-cluster variance and forms clusters by merging the two clusters that result in the smallest increase in the total within-cluster sum of squares (WCSS). This method generally produces clusters that are nearly equal in size and spherical in shape. However, for data with complex cluster shapes, the results may not be optimal (Dogan & Birant, 2022).

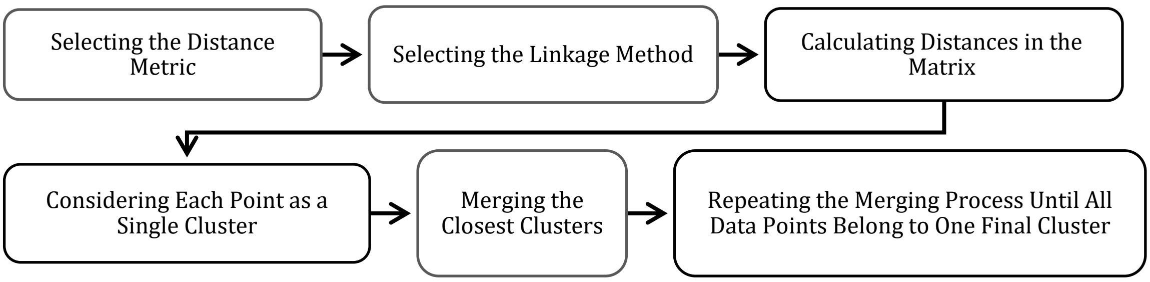
Table 1.
The methods of hierarchical clustering

Method	Distance Metric	Advantages	Disadvantages
Single Linkage	Minimum distance between closest points	Low sensitivity to cluster shape	Prone to chain-like clusters
Average Linkage	Average distance between all point pairs	Less prone to chain-like clusters, spherical clusters	Sensitive to cluster size
Ward's Method	Increase in WCSS upon merging clusters	Produces compact, spherical clusters of nearly equal size	Sensitive to complex cluster shapes

(Source: The Researcher's Findings)

In Fig. 1, the steps of the hierarchical clustering algorithm, considering the linkage method, are shown step by step. This algorithm is capable of uncovering the hierarchical structure of the data and progressively merging clusters so that one can observe and analyze the clusters more comprehensively at each stage (Asawa, 2021).

Figure 1.
Hierarchical Clustering Model Framework.



(Source: The Researcher's Findings)

6. PERFORMANCE EVALUATION METRICS

6.1 Sharpe Ratio

The Sharpe ratio, also known as the reward-to-volatility ratio, is a performance evaluation metric that measures the excess return over the risk-free rate, relative to the risk assumed by the investor (Sharpe, 1966). This metric was introduced by William Sharpe and is used to compare different investment portfolios. Generally, the Sharpe ratio is used to measure the risk of a portfolio, based on its total return (portfolio return) and the return of a risk-free investment (typically the return on government treasury bills). The Sharpe ratio shows how much additional return a portfolio has earned for each unit of risk. In comparing investment portfolios, the portfolio with a higher Sharpe ratio is usually considered superior because it indicates higher returns relative to the risk taken (Lim & Ong, 2020; Ferri, 2010).

6.2 Sortino Ratio

The Sortino ratio focuses on the downside risk of an investment portfolio and serves as an alternative to the standard deviation in the Sharpe ratio. This ratio measures the excess return for each unit of downside risk and is calculated by dividing the difference between the portfolio return and the risk-free rate by the downside deviation of the portfolio, which reflects the amount of downside risk (Venugopal & Sophia, 2020).

6.3 Maximum Drawdown

Maximum Drawdown (MDD) effectively presents the worst-case scenario for investors

(Raffinot, 2017). This metric answers the question of how much loss an investor might incur if they buy at the highest price and sell at the lowest possible price. Since this indicator measures the largest single drop from the peak to the lowest value of a portfolio's value, it is considered a downside risk metric. The Maximum Drawdown is expressed as a percentage and indicates how much of a portfolio's value has been lost relative to its highest point over a given time period. A portfolio with a lower maximum drawdown may be considered less risky and more stable. This metric shows how much an investment portfolio has lost in value during a market downturn. Generally, portfolios with lower maximum drawdowns are considered less risky during market declines because their value reduction during crisis times is relatively smaller.

6.4 Value at Risk

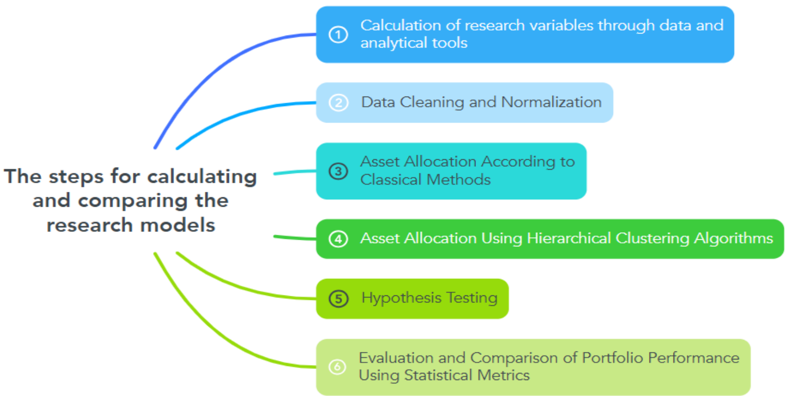
Value at Risk (VaR) is an important metric for assessing investment risk. It shows the probability that an investment might lose a certain amount of its value within a specified time period. Typically, VaR is determined at a specific confidence level, often expressed as a percentage. By calculating and monitoring VaR, investors can identify potential risks and make informed decisions about managing risk and adjusting their investment portfolios. This metric is popular among treasury managers, portfolio managers, and financial institutions, as it helps to assess how bad a situation could become. However, VaR has limitations and weaknesses and should be combined with other risk assessment tools and metrics for a more comprehensive view of investment risk (Seyfi et al., 2021).

6.5 Expected Shortfall

While VaR answers the question, "How bad can the situation get?" the Expected Shortfall (ES) asks, "What is the expected loss if the bad situation occurs?" In fact, Expected Shortfall builds upon Value at Risk. This metric is a function of the time horizon in days and the level of confidence. The Expected Shortfall is the average loss that exceeds the VaR threshold over a period of N days, assuming the loss exceeds the Xth percentile (the confidence level) of the loss distribution. This metric encourages diversification (Hallin & Trucíos, 2023).

In order to compare between different clustering methods and classical approaches, the steps are executed as shown in the figure below:

Figure 2.
Steps for Calculation and Comparison of Research Models.



(Source: The Researcher's Findings)

Literature Review

Past With the development of sophisticated algorithms that provide improved risk management techniques, the goal of creating an efficient portfolio has become a central focus of financial engineering research. One innovative approach in this domain is the Hierarchical Equal Risk Contribution (HERC) algorithm, which has gained traction for its ability to optimize portfolios by ensuring balanced risk distribution among assets. This literature review synthesizes existing research on the HERC algorithm and its application in constructing efficient portfolios. Markowitz's groundbreaking work, which focused on the trade-off between return and risk, established contemporary portfolio theory and served as the basis for portfolio optimization (Markowitz, 1952). Markowitz advocated for diversification to optimize asset allocation, a principle that still underpins many contemporary approaches. However, as financial markets have evolved, traditional methods have often been criticized for their oversimplified assumptions about asset correlations and risk distribution (Nourahmadi & Sadeqi, 2022). The application of data-driven and machine-learning-based recommender systems has demonstrated promising results in assisting investors to outperform the market by uncovering latent relationships among stocks (Nourahmadi, Rahimi, & Sadeqi, 2024). This calls for investigating alternate approaches that can more successfully deal with these constraints. The HERC algorithm, initially proposed by Raffinot (2018), leverages a hierarchical clustering framework to distribute risk equally among assets in a portfolio. This innovative approach addresses the problem of risk concentration, which is common in traditional portfolio constructions. By employing a clustering strategy, the HERC algorithm effectively identifies asset groupings that enhance diversification and stabilize returns (Duarte & De Castro, 2020). Huang (2021) and other recent researchers have assessed the HERC algorithm's performance across a range of marketplaces, proving its effectiveness in settings like the Chinese stock market. Huang's findings support the argument that HERC portfolios can yield better risk-adjusted returns compared to traditional models, thereby solidifying the importance of this algorithm in contemporary portfolio management (Huang & Gao, 2021). Furthermore, researching hierarchical clustering methods has highlighted their potential in improving the robustness of financial models under volatile market conditions (Nourahmadi & Sadeqi, 2023, Nourahmadi, et al, 2021). In conjunction with HERC, adaptive methods like adaptive seriational risk parity have emerged, which utilize machine learning techniques to optimize portfolio construction dynamically. The extensions of these adaptive techniques were examined by (Schwendner et al., 2021), who demonstrated how incorporating machine learning might improve risk parity strategies' performance in intricate financial contexts. These advancements indicate a shifting paradigm where machine learning tools are being harnessed alongside traditional financial theories to create more proactive portfolio management solutions. The interplay of hierarchical clustering and risk parity has also been emphasized in the works of (Millea & Edalat, 2022) and (Deković & Šimović, 2025). Millea and Edalat specifically focus on using deep reinforcement learning in conjunction with hierarchical risk parity, further illustrating the innovative combinations of techniques that can yield optimal portfolio compositions. Similarly, Deković and Šimović's research provides insights into the efficient implementation of hierarchical risk strategies in real-world settings, underscoring the significant implications for practitioners and portfolio managers. Moreover, research conducted by Sen et al (Sen & Mehtab, 2021; Sen & Dutta, 2023) on risk-based portfolio optimization in specific market sectors corroborates the relevance of HERC by suggesting that effective risk distribution not only enhances stability but also tailors investment strategies to better fit market dynamics. Their comparative studies highlight the necessity of incorporating advanced algorithms like HERC to develop portfolios that can adapt to the unique behaviors of various asset classes. The application of clustering methods in finance has further been explored by Cajas (Cajas, 2023), who employed graph theory to facilitate portfolio optimization. This research aligns with the findings of Nourahmadi & Sadeqi (Nourahmadi & Sadeqi, 2021), reinforcing the idea that clustering techniques can play a vital role in enhancing portfolio diversification processes. The synergistic effects of these methodologies open doors for further exploration into how these algorithms can be integrated to improve investment outcomes significantly. In summary, the literature indicates a growing consensus on the effectiveness of the hierarchical equal risk contribution algorithm as a pivotal tool in achieving portfolio efficiency. Its capacity to balance risk distribution among assets, coupled with the utilization of modern technologies such as machine learning and clustering techniques, illustrates the evolution of risk management in portfolio construction. As markets become increasingly complex, embracing methodologies like HERC will be crucial in developing strategies that not only safeguard against inherent financial risks but also aim for optimal returns across diversified portfolios. Future research should continue to investigate the integration of HERC with other computational techniques to further enhance portfolio performance and address the nuances of various market conditions.

Research Hypotheses

Hypothesis 1: The performance of the Hierarchical Equal Risk Contribution (HERC) algorithm is superior to that of the Equal Weight (EW) portfolio model.

Hypothesis 2: The performance of the Hierarchical Risk Parity (HRP) algorithm is superior to that of the Inverse Variance (IV) model.

Methodology

This study adopts a quantitative approach and employs numerical and statistical methods to identify and evaluate the optimal model. From an applied perspective, the objective is to develop a stock asset allocation framework that maximizes returns while minimizing risk, thereby providing practical insights for asset managers as well as individual and institutional investors. From a methodological standpoint, the research is conducted within a positivist paradigm and follows a deductive approach, whereby hypotheses derived from the existing theoretical literature on portfolio optimization and asset allocation algorithms are empirically tested using capital market data. The research strategy is based on correlation analysis and the comparison of models under identical conditions, with the aim of assessing the relative performance of different asset allocation methods. Data collection combines library research with the extraction of real market data, which form the basis for the quantitative analyses and statistical tests employed in the study. Initially, the adjusted stock prices of the sample (accounting for dividend payments and capital increases) over a ten-year period were extracted. After normalizing and cleansing the data, the second stage involved calculating classical models, including the Hierarchical Equal Risk Contribution (HERC) model, Hierarchical Risk Parity (HRP) algorithm, Inverse Variance (IV) model, and the Equal Weight (EW) portfolio, using Python software under the study's assumptions. In the third stage, the profitability and risk performance of these models were assessed using modern and post-modern portfolio metrics. Finally, the fourth stage was dedicated to comparing and evaluating the superiority of each strategy's performance through statistical tests.

Population and Sample

The statistical population comprises securities of all active issuers on the Tehran Stock Exchange (TSE) and Iran Farabourse market during the period from the beginning of 2013 to the end of 2022. The sample consists of daily data from 88 companies listed on the TSE that meet the following criteria:

1. The companies must have been listed on the TSE by the end of 2022.
2. Their stock symbols must have been active on the trading board throughout the study period with a complete price history.
3. The stocks must have been traded on at least 160 business days per year.

Sampling covered the period from 03/21/2013 to 03/18/2023, encompassing 2,408 trading days. Data were collected using Excel and subsequently processed, cleaned, and analyzed. Algorithms were executed, variables computed, and hypotheses tested using Python version 3.9.7 and libraries including NumPy, Pandas, Matplotlib, Seaborn, SciPy, Scikit-Learn, FinPy, and CVXPY. Statistical tests were performed in SPSS, while Excel was used for data visualization and chart creation.

Findings

Table 2 provides a summary of the descriptive statistics for the returns of the various portfolio strategies analyzed in the study. These statistics offer a comparative overview of the central tendency, dispersion, and risk-return characteristics of each strategy.

A. Statistical Analysis

The statistical analysis shows that the HERC single machine learning algorithm generally outperforms other machine learning models and classical methods such as EW and IV in terms of mean return. This indicates the potential of the HERC single algorithm to achieve higher returns. However, this algorithm also exhibits moderate volatility and higher kurtosis, suggesting heavier tails in the distribution of returns. HRP average has the highest standard deviation (0.199), indicating the most significant fluctuations. Both HRP single and HRP average demonstrate high kurtosis, reflecting heavy-tailed distributions and a higher likelihood of extreme positive or negative returns, representing higher risk for these strategies.

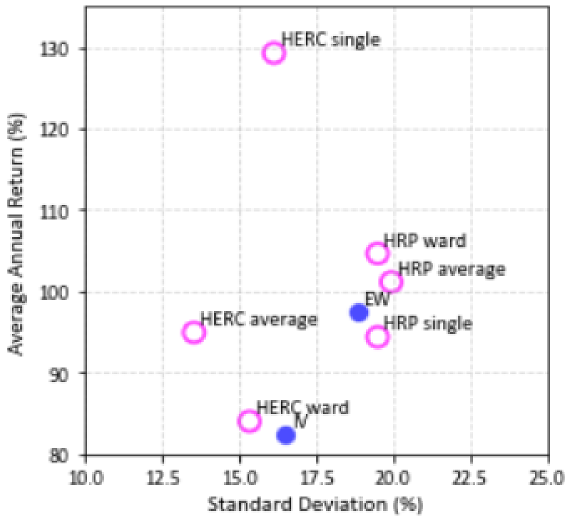
Table 2.
Central Tendency and Dispersion Indicators of Portfolio Returns

Algorithm	Mean	Standard Deviation	Min	25 Percentile	Median	75 Percentile	Max	Skewness	Kurtosis
HRP ward	1.046	0.195	-0.346	0.066	0.619	1.575	4.781	1.927	4.146
HRP single	0.944	0.195	-0.357	0.061	0.57	1.313	4.691	2.144	5.308
HRP average	1.012	0.199	-0.397	-0.009	0.569	1.441	5.241	2.165	5.336
HERC ward	0.84	0.154	-0.351	0.013	0.482	1.321	4.158	2.067	5.053
HERC single	1.293	0.162	-0.25	0.044	0.585	1.289	8.328	2.888	8.725
HERC average	0.949	0.136	-0.353	0.062	0.477	1.417	4.648	2.099	5.193
IV	0.824	0.165	-0.297	0.019	0.491	1.187	4.082	2.092	5.116
EW	0.976	0.188	-0.332	0.054	0.598	1.379	4.641	2.028	4.662

(Source: The Researcher's Findings)

In contrast, classical algorithms like EW and IV exhibit greater stability: EW, with a mean return of 0.976 and a standard deviation of 0.188, provides a reasonable balance between return and volatility. IV demonstrates better performance in adverse market conditions, with the smallest negative minimum return (-0.297), highlighting its resilience in critical scenarios.

Figure 3.
Return versus Risk Visualization for Different Portfolios.

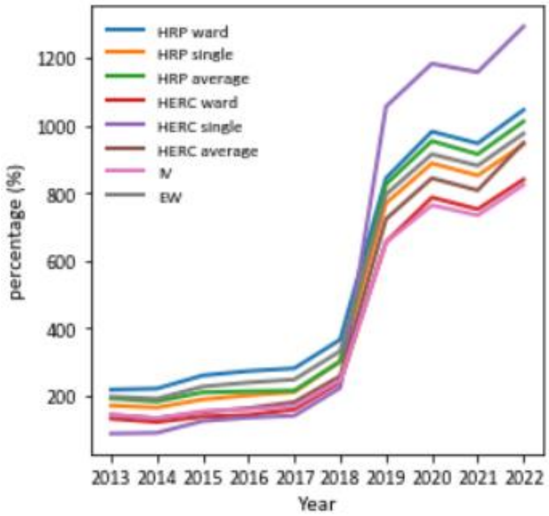


(Source: The Researcher's Findings)

The analysis of the chart indicates that the HERC single strategy achieves a better balance between return and risk compared to other methods. This approach allows investors to pursue their financial goals with greater confidence. Regarding the HERC methods, it is observed that their standard deviation is lower than that of the classical methods, EW and IV, which demonstrates the superiority of these methods over the two classical approaches, EW and IV.

Based on the data presented in Fig 4, the HERC single strategy significantly outperforms the other strategies, while the IV strategy recorded the lowest returns.

Figure 4.
Cumulative Return Trends of Different Portfolios.



(Source: The Researcher's Findings)

B. Risk-Based Performance Evaluation Variables

In this section, the performance of each classic and machine learning portfolio has been evaluated using the metrics presented in Table 3, which include maximum drawdown, annual VaR (5%), annual CVaR (5%), Sharpe ratio, and Sortino ratio. It highlights the superior performance of the HERC Single strategy in terms of risk-adjusted returns, evidenced by its high Sharpe and Sortino ratios.

Table 3.
Average Risk-Based Performance Metrics of Classical and Machine Learning Portfolios

Algorithm	MD	VaR	CVaR	Sharpe	Sortino
Machine Learning					
HRP Ward	25.5%	29.8%	37.1%	2.38	4.31
HRP Single	26.5%	30.9%	38.3%	2.13	3.78
HRP Average	27.9%	30.8%	39.2%	2.14	3.84
HERC Ward	20.7%	23.1%	30.5%	2.01	3.67
HERC Single	19.4%	23.0%	31.2%	4.62	6.89
HERC Average	17.4%	19.5%	25.7%	2.74	5.38
Classical					
IV	22.4%	25.8%	32.88%	2.14	3.84
EW	25.4%	29.1%	36.82%	1.98	3.75

(Source: The Researcher's Findings)

In terms of Maximum Drawdown, machine learning algorithms, particularly HRP Average with a drawdown of 27.96%, exhibit the highest maximum loss, indicating an increased likelihood of negative returns during critical market conditions. Other machine learning algorithms, such as HRP Single and HRP Ward, also experience notable drawdowns. Conversely, machine learning models like HERC demonstrate lower drawdowns, reflecting more stable performance and reduced risk, especially in bearish market scenarios.

For VaR at 5% (Value at Risk at 5%), HRP Single has the highest value at 30.93%, signifying higher risk levels for this model. Other machine learning models, such as HRP Average and HRP Ward, also show varying levels of risk exposure.

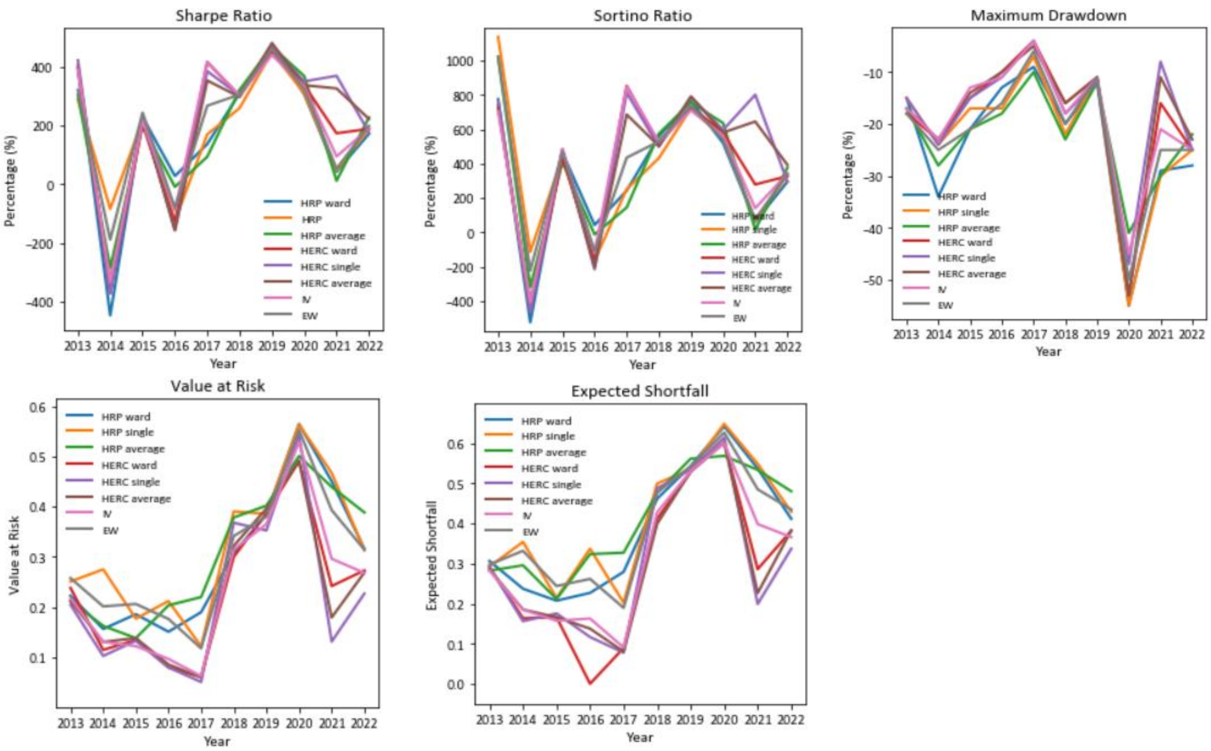
Regarding CVaR at 5% (Conditional Value at Risk at 5%), HRP Average records the highest value at 39.24% among machine learning algorithms, indicating a higher risk tolerance. Classical models such as EW (36.82%) and IV (32.88%) exhibit greater risk compared to the HERC machine learning models.

In terms of the Mean Sharpe Ratio, machine learning algorithms, including HERC Single (4.62) and HERC Average (2.74), achieve the highest Sharpe ratios, indicating favorable returns relative to risk. Classical models such as IV (2.14) also perform reasonably well, but machine learning algorithms outperform them on this metric. Finally, for the Mean Sortino Ratio, HERC Single (6.89) and HERC Average (5.38) demonstrate the highest values, reflecting their superior ability to manage downside risk. Classical models like IV (3.84) and EW (3.75) offer similar results for this ratio, but machine learning algorithms, particularly HERC Single, show significantly better performance in this aspect.

In conclusion, while machine learning algorithms provide varying results on risk measures such as VaR and CVaR, they achieve higher Sharpe and Sortino ratios compared to classical models. Classical models like EW and IV are more stable and exhibit lower risk, but in terms of risk-adjusted returns, machine learning algorithms, especially HERC Single and HERC Average, demonstrate superior performance.

Figure 5.

Performance metrics of various portfolios from 21-3-2013 to 19-3-2023 - from top left to bottom: Sharpe ratio, Sortino ratio, Maximum Drawdown, Value at Risk, Expected Shortfall.



(Source: The Researcher's Findings)

C. Hypothesis Testing of the Research

The Kruskal-Wallis test is a non-parametric statistical method used to compare the rankings of different groups, particularly when there are two or more groups involved. This test provides two main outputs (Safavi Iranji et al., 2024).

1. Rankings of Groups:
In the first output, groups are ranked based on the variables under consideration. Higher rankings indicate better performance of the method for a given metric, while lower rankings suggest weaker performance for the same metric.

2. Test Statistics and Significance Levels:
The second output reports the Kruskal-Wallis test statistics and their corresponding significance levels. In this study, hypotheses are considered supported if the significance level is below 10%.

There is a direct relationship between higher values and rankings of metrics like the Sharpe and Sortino ratios and the improved performance of models. In other words, models achieving higher values and rankings in these metrics demonstrate superior risk-adjusted returns and a better return-to-volatility ratio.

Conversely, an inverse relationship is observed between lower values and rankings of metrics like Maximum Drawdown, Value at Risk (VaR), and Conditional Value at Risk (CVaR) with better model performance. This means that models with lower values and rankings in these criteria exhibit greater stability and reduced risk of loss.

The results of the Kruskal-Wallis test rankings, based on the five selected criteria and the clustering linkage methods, are presented in Table 4.

Table 4.
Kruskal-Wallis Statistical Test Results

Metric	Mean Rank		KW	Mean Rank		KW	Mean Rank		KW
	EW	HERC single		EW	HERC average		EW	HERC ward	
Sharpe Ratio	7.8	13.4	3.221**	9.7	11.3	0.366	10.4	10.6	0.006
Sortino Ratio	7.85	13.55	3.014**	9.5	11.5	0.571	10.6	10.4	0.006
Maximum Drawdown	12.5	8.5	2.286	12.5	8.5	2.286	11.9	9.1	1.12
Conditional VaR (CVaR)	11.6	9.4	0.691	13	8	3.571**	11.9	9.1	1.12
Value at Risk (VaR)	12.1	8.85	1.557	13	8	3.571**	12	9	1.286
Metric	**IV**	**HRP single**	**KW**	**IV**	**HRP average**	**KW**	**IV**	**HRP ward**	**KW**
Sharpe Ratio	10.6	10.4	0.006	10.8	10.2	0.051	10.2	10.8	0.051
Sortino Ratio	10.4	10.6	0.006	10.7	10.3	0.023	10.2	10.8	0.051
Maximum Drawdown	9.5	11.5	0.571	9.2	11.8	0.966	9.7	11.3	0.366
Conditional VaR (CVaR)	9	12	1.286	8.8	12.2	1.651	9.4	11.6	0.691
Value at Risk (VaR)	9	12	1.286	9	12	1.286	9.1	11.9	1.12

Notes:
KW indicates the Kruskal-Wallis test statistic.
** indicate significance at the 10% level.
(Source: The Researcher's Findings)

The analysis of the Kruskal-Wallis test results indicates the following:

Sharpe Ratio: A significant difference at the 10% level was found between EW and HERC Single, indicating the superior performance of HERC Single in this metric. However, no significant differences were observed between EW and HERC Average, as well as EW and HERC Ward, suggesting that the performance of these methods is similar in this metric. Sortino Ratio: A significant difference at the 10% level was observed between EW and HERC Single, with HERC Single performing better. However, when comparing EW with HERC Average and EW with HERC Ward, no significant difference was found, indicating comparable performance between these methods. Maximum Drawdown: No significant differences were found between EW and the HERC models (Single, Average, Ward) in terms of maximum drawdown. This suggests that these models exhibit similar performance in minimizing losses during market downturns. Conditional Value at Risk (CVaR): No significant difference was found between EW and HERC Single. However, HERC Average showed better performance than EW at the 10% level, while no significant difference was reported between EW and HERC Ward. Value at Risk (VaR): No significant

differences were found in comparisons of EW with the HERC Single and HERC Ward models. The IV and HRP models (Single, Average, Ward) generally showed similar performance across most metrics. Although in some cases, such as with the Sharpe and Sortino ratios, the KW value approached significance, no statistical significance was reported at the 0.1 level.

The results suggest that HERC strategies outperform traditional strategies such as EW and IV in certain metrics, particularly in risk-adjusted return metrics like the Sharpe and Sortino ratios. These findings indicate that using data-driven and machine learning-based models can effectively improve portfolio management and lead to higher returns.

Discussion and Conclusion

The results of this research indicate that the Hierarchical Equal Risk Contribution (HERC) algorithm, as an innovative tool, shows remarkable potential in portfolio management and transforming risk into opportunity in financial markets. The HERC models, especially the single and average versions, exhibited superior performance in risk-adjusted return metrics, such as the Sharpe ratio and Sortino ratio, compared to traditional methods like Equal Weighting (EW) and Inverse Variance (IV). These findings demonstrate the ability of HERC to create an efficient balance between returns and risk management.

At the same time, the results showed that there was no significant difference in the maximum drawdown metric between the methods. These results, along with the superior performance of HERC, suggest that the algorithm is capable of maintaining stability and providing a conservative approach to risk management. Therefore, HERC can be considered as an optimal option for investors seeking to reduce risk and enhance returns. By leveraging clustering structures and analyzing nonlinear relationships between assets, HERC has secured a unique position among modern portfolio management methods. Additionally, by integrating hierarchical approaches and risk analysis, this algorithm goes beyond traditional methods and provides potential capacities for improving investment decision-making. However, to fully exploit this method, it is recommended that its performance be evaluated under various market conditions and economic cycles.

The findings of this research align with previous studies, such as the work of Vito Ciciretti (Ciciretti & Bucci, 2023), which showed that the use of clustering methods for portfolio optimization outperforms classical approaches. Furthermore, external studies, such as the research by Thomas Raffinot (2018), also confirm the results of this study. Raffinot's research showed that HERC-based portfolios, using a hierarchical approach and equal risk contribution allocation, performed better than other methods in metrics such as Conditional Drawdown at Risk (CDaR). Additionally, the findings of Lopez de Prado (2016) demonstrated that the HRP method (the parent approach of HERC) resolves instability issues and focuses on the problems of Markowitz methods, producing more diversified portfolios with lower risk.

Despite the positive findings, some studies, like Jain and Jain (2019), have shown that hierarchical algorithms, including HERC and HRP, are less sensitive to incorrect covariance estimates compared to methods like minimum variance or maximum diversification. This emphasizes the importance of accurate input parameter estimation to fully optimize the performance of HERC. To complement these findings and take further advantage of HERC's potential, it is recommended that the algorithm's performance be evaluated in various market conditions, as assessing its compatibility in volatile periods and different economic cycles can provide a better understanding of its stability. Additionally, combining HERC with classical methods could create a hybrid approach that leverages the strengths of both data-driven and traditional methods, leading to improved portfolio performance. Moreover, refining and improving input parameters and analyzing the impact of nonlinear relationships in financial data could contribute to further development of this algorithm.

Finally, this study, by highlighting the advantages of the HERC algorithm, demonstrates that this method can bring a significant transformation to portfolio management and serve as an efficient alternative to traditional methods. HERC not only offers higher returns but also, with its innovative risk management approach, provides investors with a powerful tool to turn risk challenges into investment opportunities.

Abbreviations

HERC ward	Hierarchical Equal Risk Contribution using Ward's Method Linkage
HERC single	Hierarchical Equal Risk Contribution Single Method Linkage
HERC average	Hierarchical Equal Risk Contribution using Average Method Linkage
HRP ward	Hierarchical Risk Parity using Ward's Method Linkage
HRP single	Hierarchical Risk Parity using Single Method Linkage
HRP average	Hierarchical Risk Parity using Average Method Linkage
IV	Inverse Variance
EW	Equal Weight Portfolio
Sharpe	Sharpe Ratio
Sortino	Sortino Ratio
MD	Maximum Drawdown
VaR	Value at Risk at 5.0%
CVaR	Conditional Value at Risk
KW	Kruskal-Wallis test statistic

REFERENCE

Asawa, Y. (2021). Modern machine learning solutions for portfolio selection. IEEE Engineering Management Review, 50(1), 94–112. https://ieeexplore.ieee.org/abstract/document/9627810.

Cajas, D. (2023). A graph theory approach to portfolio optimization. SSRN Working Paper. https://dx.doi.org/10.2139/ssrn.4602019.

Ciciretti, V., & Bucci, A. (2023). Building optimal regime-switching portfolios. Economic Modelling and Finance, 64, S1062940822001723. https://doi.org/10.1016/j.najef.2022.101837.

Deković, D., & Šimović, P. P. (2025). Hierarchical risk parity: Efficient implementation and real world analysis. Future Generation Computer Systems, 107, 107744. https://doi.org/10.1016/j.future.2025.107744.

DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? Review of Financial Studies, 22(5), 1915–1953. https://doi.org/10.1093/rfs/hhm075.

Dogan, A., & Birant, D. (2022). K-centroid link: A novel hierarchical clustering linkage method. Applied Intelligence, 52, 1–24. https://doi.org/10.1007/s10489-021-02624-8.

Duarte, F. G., & De Castro, L. N. (2020). A framework to perform asset allocation based on partitional clustering. IEEE Access, 8, 110775–110788. https://doi.org/10.1109/access.2020.3001944.

Ferretti, S. (2022). On the modeling and simulation of portfolio allocation schemes: An approach based on network community detection. Computational Economics, 62(3), 969–1005. https://link.springer.com/article/10.1007/s10614-022-10288-w.

Ferri, R. (2010). All about asset allocation (2nd ed.). McGraw-Hill Education. https://search.worldcat.org/title/663882714.

Hallin, M., & Trucíos, C. (2023). Forecasting value-at-risk and expected shortfall in large portfolios: A general dynamic factor model approach. Econometrics and Statistics, 27, 1–5. https://doi.org/10.1016/j.ecosta.2021.04.006.

Huang, W., & Gao, X. (2021). Evaluating hierarchical equal risk contribution portfolios in the Chinese stock market. Journal of Mathematical Finance, 12(1), 179–195. https://doi.org/10.4236/jmf.2022.121011.

Jain, P., & Jain, S. (2019). Can machine learning-based portfolios outperform traditional risk-based portfolios? The need to account for covariance misspecification. Risks, 7(3), 74. https://doi.org/10.3390/risks7030074.

Lim, T., & Ong, C. (2020). Portfolio diversification using shape-based clustering. Journal of Financial Data Science, 3(1), 111. https://doi.org/10.3905/jfds.2020.1.054.

López de Prado, M. (2016). Building diversified portfolios that outperform out of sample. Journal of Portfolio Management, 42(4), 59–69. https://doi.org/10.3905/jpm.2016.42.4.059.

Maillard, S., Roncalli, T., & Teïletche, J. (2010). The properties of equally weighted risk contribution portfolios. Journal of Portfolio Management, 36(4), 60–70. https://doi.org/10.3905/jpm.2010.36.4.060.

Markowitz, H. M. (1952). Portfolio selection. Journal of Finance, 7(1), 77–91. https://doi.org/10.2307/2975974.

Menvouta, E. J., Serneels, S., & Verdonck, T. (2023). Portfolio optimization using cellwise robust association measures and clustering methods with application to highly volatile markets.

Journal of Financial Data Science, 9, 100097. https://doi.org/10.1016/j.jfds.2023.100097.

Millea, A., & Edalat, A. (2022). Using deep reinforcement learning with hierarchical risk parity for portfolio optimization. International Journal of Financial Studies, 11(1), 10. https://doi.org/10.3390/ijfs11010010.

Nourahmadi, M., & Sadeqi, H. (2021). Hierarchical risk parity as an alternative to conventional methods of portfolio optimization: A study of Tehran Stock Exchange. Iranian Journal of Finance, 5(4), 1–24. https://doi.org/10.30699/ijf.2021.289848.1242.

Nourahmadi, M., & Sadeqi, H. (2022). A machine learning-based hierarchical risk parity approach: A case study of a portfolio consisting of stocks of the top 30 companies on the Tehran Stock Exchange. Financial Research Journal, 24(2), 236–256. https://doi.org/10.22059/frj.2021.319092.1007146.

Nourahmadi, M., & Sadeqi, H. (2023). Portfolio diversification based on clustering analysis. Iranian Journal of Accounting, Auditing and Finance, 7(3), 1–16. https://doi.org/10.22067/ijaaf.2023.43078.1092.

Nourahmadi, Marziyeh, rasti, Fatemeh, & SADEGHI, HOJATOLLAH. (2021). A Review of Research on Financial Time Series Clustering: A Bibliometrics Approach. ADVANCES IS FINANCE AND INVESTMENT, 2(2), 23-57. SID. https://sid.ir/paper/391770/en

Nourahmadi, M. , Rahimi, A. and Sadeqi, H. (2024). Designing a Stock Recommender System Using the Collaborative Filtering Algorithm for the Tehran Stock Exchange. *Financial Research Journal*, *26*(2), 318-346. doi: 10.22059/frj.2023.360955.1007479

Raffinot, T. (2017). Hierarchical clustering-based asset allocation. Journal of Portfolio Management, 44(2), 89–102. https://doi.org/10.3905/jpm.2018.44.2.089.

Raffinot, T. (2018). The hierarchical equal risk contribution portfolio. SSRN Electronic Journal. https://dx.doi.org/10.2139/ssrn.3237540.

Rostami, M. , Rasti, F. and Abbasi, E. (2025). Copula-Based Risk Modeling: A Comparative Analysis of MCAViaR and Gaussian Copulas for Global Indices. *Journal of Mathematics and Modeling in Finance*, *5*(2), 77-106. doi: 10.22054/jmmf.2025.86227.1187

Safavi Iranji, M., Zanjirdar, M., Safa, M., & Jahangirnia, H. (2024). Asset allocation using nested clustered optimization algorithm: A novel approach to risk management in portfolio. Journal of Mathematical Modeling in Finance, 4(2), 137–157. https://doi.org/10.22054/jmmf.2025.82388.1149.

Sajadi, S. M. A., Barak, S., & Fereydooni, A. (2024). Online portfolio selection using macroeconomic pattern matching. SSRN Electronic Journal. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4960922.

Schwendner, P., Papenbrock, J., Jaeger, M., & Krügel, S. (2021). Adaptive seriational risk parity and other extensions for heuristic portfolio construction using machine learning and graph theory. Journal of Financial Data Science, 3(4), 65–83. DOI: 10.3905/jfds.2021.1.078.

Sen, J., & Dutta, A. (2023). Portfolio optimization for the Indian stock market. In Encyclopedia of Data Science and Machine Learning (pp. 1904–1951). IGI Global. https://www.igi-global.com/chapter/portfolio-optimization-for-the-indian-stock-market/317595.

Sen, J., & Mehtab, S. (2021). A comparative study of optimum risk portfolio and eigen portfolio on the Indian stock market. International Journal of Business Forecasting and Marketing Intelligence, 7(2), 143–193. https://doi.org/10.1504/IJBFMI.2021.120155.

Seyfi, S. M., Sharifi, A., & Arian, H. (2021). Portfolio value-at-risk and expected shortfall using an efficient simulation approach based on Gaussian mixture model. Mathematics and

Computers in Simulation, 190, 1056–1079. https://doi.org/10.1016/j.matcom.2021.05.029.

Sharpe, W. F. (1966). Mutual fund performance. Journal of Business, 39(1), 119–138. https://www.jstor.org/stable/2351741?searchText=Mutual%20fund%20performance&searchUri=%2Faction%2FdoBasicSearch%3FQuery%3DMutual%2Bfund%2Bperformance%26so%3Drel&ab_segments=0%2Fbasic_search_gsv2%2Fcontrol&refreqid=fastly-default%3A32de7d5ea9a6a553b1e108b638c1f88f.

Sjöstrand, D., Behnejad, N., & Richter, M. (2020). Exploration of hierarchical clustering in long-only risk-based portfolio optimization (Doctoral dissertation). Copenhagen Business School, Copenhagen. https://research.cbs.dk/files/62178444/879726_Master_Thesis_Nima_Daniel_15736.pdf.

Venugopal, M., & Sophia, S. (2020). Examining Sharpe ratio, ASR, Sortino, Treynor and information ratio in Indian equity mutual funds during the pandemic. International Journal of Management, 11(11). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3780440.

Zhao, Y., & Karypis, G. (2002). Evaluation of hierarchical clustering algorithms for document datasets. In Proceedings of the International Conference on Information and Knowledge Management (pp. 515–524). https://doi.org/10.1145/584792.584877.

Journal of
Knowledge Economy Studies

https://kes.hmu.ac.ir

Hazrat-e Masoumeh University

When Sustainability Meets Machine Learning: Reinforcement and Neural Evidence from an Emerging Market

Article Type:
Research Article

Setareh Azadvar[1]
Department of accounting, Shandiz Institute of Higher Education, Mashhad, Iran
E-mail: setareh.azadvar@gmail.com

Alireza Azarberahman[2*]
Department of Accounting, Shandiz Institute of Higher Education, Mashhad, Iran
Corresponding Author
E-mail: a.azarberahman@shandiz.ac.ir

Received: 14 November 2025
Revised: 15 December 2025
Accepted: 17 March 2026
Available Online: 18 March 2026

Abstract

This study examined how firm-level environmental and social performance relates to stock price volatility in an emerging market characterized by limited transparency and weaker institutional frameworks. While prior research largely relied on linear models and focused on developed economies, this study adopted a dynamic, data-driven perspective to capture potentially nonlinear and time-dependent sustainability–risk patterns. Using a panel of non-financial firms listed on Tehran Stock Exchange (TSE) over the period 2011–2023, the firm-level environmental and social indicators were constructed based on a systematic analysis of sustainability disclosures. Empirical results from conventional linear regressions indicated weak and statistically insignificant average associations between sustainability performance and stock volatility. However, learning-based models, including reinforcement learning (RL) and LSTM neural networks, demonstrated superior ability to capture nonlinear and dynamic volatility patterns conditional on sustainability-related information. These findings suggested that the sustainability disclosures contain predictive information for volatility dynamics, even when linear risk-reduction effects are not evident. The study highlighted the importance of flexible modeling frameworks when assessing the financial implications of environmental and social performance in emerging markets.

Keywords

ESG performance, Stock volatility, Reinforcement learning (RL), LSTM neural networks, Sustainable finance, Nonlinear dynamics.

Cite this article: Azadvar, S. & Azarberahman, A., (2026). When sustainability meets machine learning: reinforcement and neural evidence from an emerging market. *Journal of Knowledge Economy Studies (JKES)*, 3(1), 47-71. DOI: http//doi.org/10.22034/kes.2026.2077788.1087

Online ISSN: 3060-7329

Introduction

Corporate sustainability has migrated from being a discretionary add-on into a fundamental strategic dimension that shapes firm value, risk exposure, and stakeholder trust (Naseer et al., 2024). Evidence from developed markets suggests that higher environmental, social, and governance (ESG) performance is often associated with lower equity risk and cost of capital (Biju, 2025; Gupta & Chaudhary, 2023). However, these findings are derived from institutional environments characterized by strong disclosure regimes and investor protection, and therefore cannot be readily generalized to emerging markets (Biju, 2025; Gupta & Chaudhary, 2023). Yet in emerging markets, where institutional frameworks are weaker, transparency is limited, and standardized ESG ratings are sparse, the empirical evidence remains fragmented (Kansoy & Stasiulaitis, 2025; Rahat & Nguyen, 2024).

Environmental and social initiatives can reduce firm-specific risks by enhancing operational resilience, stakeholder relationships, and regulatory compliance (Iannone et al., 2025). However, such practices may also introduce new uncertainties and transitional cost burdens, particularly in less-developed institutional settings (Farūq & Chowdhury, 2025). The dual nature of the sustainability's impact on market risk — reduction through improved governance vs. increase through complexity and uncertainty — presents a substantive empirical puzzle. In particular, in a market context like Iran's, where ESG disclosures are largely voluntary and the investor protection is still evolving, assessing how sustainable practices influence the formation of equity market risk is critically important.

Prior research in emerging economies has primarily examined ESG performance's relation to firm valuation or profitability, often using static linear models and binary disclosure proxies. For example, Rahat and Nguyen (2024) reported on the impact of ESG on firm valuation in emerging markets but they did not focus on risk dynamics. Biju (2025) assessed the ESG-firm performance nexus and acknowledged that the direct link to firm risks remains underexplored. Meanwhile, normative frameworks in developed contexts do not necessarily transfer to emerging-market realities (Valencia Söderberg & Truong, 2024). Examining non-linear, temporally-dependent relationships between firm-level sustainability practices and equity risk in emerging markets, using advanced analytical methods suited to dynamic, state-contingent environments is a gap in literature. Importantly, addressing this gap requires moving beyond average linear effects toward models that prioritize risk dynamics and predictive performance, rather than causal inference based on static specifications. This study addressed these gaps in three major ways. First, it developed a context-specific, replicable environmental and social index in an Iranian context, where commercial ESG ratings are scarce. Second, it employed a hybrid empirical framework, combining econometrics, RL, and recurrent neural-network modeling, to reveal dynamic relations between sustainability performance and equity volatility. Third, it provided novel evidence from an emerging market context by examining how sustainability-related disclosures relate to the predictability and dynamics of stock volatility under conditions of weaker institutional infrastructure.

By integrating sustainability disclosure, advanced modeling, and an emerging-market setting, this research responds to the urgent call for more data-driven, risk-focused studies of corporate sustainability's financial implications (Biju, 2025; Kansoy & Stasiulaitis, 2025). It extends the frontier of knowledge on how sustainability shapes the firm-level equity risk beyond mature markets into institutional terrains with higher informational opacity and distinct regulatory challenges.

Literature Review and Hypotheses Development

Environmental Performance and Firm-level Stock Volatility

In theory, higher environmental performance should attenuate firm-level risk by reducing operational uncertainty, regulatory exposure, and reputational shocks, thereby stabilizing cash flows and dampening equity volatility. Stakeholder theory predicts that firms which meet salient stakeholder expectations (e.g., emissions abatement, resource efficiency, and compliance) enjoy greater legitimacy and smoother stakeholder relations, translating into lower perceived risk and more stable pricing in capital markets (Donaldson & Preston, 1995). Agency theory adds that credible environmental disclosure narrows information asymmetry between managers and outside investors, lowering the variance of belief updates that feed into price formation (Jensen & Meckling, 1976). Finally, signaling theory implies that costly, verifiable environmental actions serve as credible signals of managerial quality and long-run resilience. Such signals are more readily capitalized into prices when governance and disclosure quality make them believable, hence reducing volatility (Spence, 1973).

Empirically, several strands of evidence supported a volatility-reducing role of environmental (E) performance. Using a large sample from an emerging market, Liu et al. (2023) identified a causal reduction in firm-specific (idiosyncratic) volatility attributable to stronger ESG performance; their identification strategy and robustness checks point to information-environment improvements as a key channel. Complementary evidence showed that higher ESG scores lower stock price fragility—the propensity for sharp price collapses following adverse shocks—again consistent with reduced sensitivity of investors to noisy signals when sustainability practices are strong (Wang et al., 2023). Cross-market studies further documented that ESG portfolios often exhibit lower risk-adjusted variability than conventional benchmarks, with environmental quality a principal driver of the risk differential (Gupta & Chaudhary, 2023). Together, these findings indicated that environmental improvements reduce both the amplitude of day-to-day idiosyncratic fluctuations and the tail risk of crash-like episodes.

Mechanistically, three pathways link environmental performance to lower volatility. First, the operational-efficiency channel: cleaner production, energy efficiency, and waste minimization contain cost variability and supply-chain disruptions, smoothing earnings and investor expectations. Second, the regulatory-and-litigation channel: better environmental profiles lower the probability and severity of regulatory penalties and contingent liabilities, compressing the distribution of adverse cash-flow shocks and the priced risk premium (Chava, 2014). Third, the information channel: richer decision-useful environmental disclosure reduces information asymmetry and curtails noisy trading, which directly lowers idiosyncratic volatility in microstructure-based frameworks. Consistent with these channels, recent studies showed that improvements in environmental scores reduce the cost of capital, a close cousin of priced risk, reinforcing the expectation of lower observed volatility when E performance strengthens.

A possible concern is that environmental initiatives could raise short-run costs and thus increase volatility. However, equilibrium models of corporate responsibility showed that when sustainability investments enhance product differentiation and stakeholder loyalty, they lower systematic risk and improve value despite upfront costs (Albuquerque et al., 2019). In markets with credible disclosure and governance, these benefits dominate, yielding a net reduction in volatility. The empirical evidence cited above aligns with this prediction and suggests the effect extends to idiosyncratic risks in emerging markets as well.

Theory and evidence converge on a clear prediction showed that as firms raise environmental performance—through measurable, disclosed actions that stakeholders and markets can verify—firm-level stock volatility should decline via operational, regulatory, and informational channels. Hence,

> **H1:** Environmental performance is expected to be associated with lower stock volatility; however, this association may be weak or non-linear in emerging markets characterized by limited transparency.

Social Performance and Firm-level Stock Volatility

Social performance—covering employees' well-being, diversity and inclusion, community engagement, customer-centric practices, and social transparency—creates intangible capital that stabilizes firms' cash flows and investors' beliefs. From a stakeholder-theoretic view, addressing salient social claims strengthens legitimacy and trust, dampening the arrival of adverse, belief-shifting signals that would otherwise amplify the price volatility (Donaldson & Preston, 1995). Agency theory predicts that richer social disclosure narrows the information asymmetry between insiders and outside investors, reducing the variance of posterior beliefs and thus the idiosyncratic component of returns (Dhaliwal et al., 2011). A complementary "social-capital" perspective shows that investments in CSR accumulate relational assets (trust with employees, customers, and creditors), which pay off precisely in stressed states by cushioning operating and financing frictions that would translate into pronounced price swings (Lins et al., 2017).

A growing empirical literature linked stronger social practices to lower firm-level risk through multiple channels. In labor markets, high-quality employee relations proxy for superior human-capital management and lower disruption risk; firms recognized for employee satisfaction earned persistent abnormal returns and exhibited steadier fundamentals—evidence consistent with lower uncertainty premia and reduced volatility (Edmans, 2011). In information environment, initiating CSR disclosure—particularly among firms with superior social performance—reduces the cost of equity capital, a priced manifestation of risk that typically co-moves with idiosyncratic volatility (Dhaliwal et al., 2011). In crisis settings, firms with higher CSR intensity ("social capital") experienced markedly better performance, consistent with volatility dampening when shocks hit and relational buffers matter most (Lins et al., 2017). Large-sample asset-pricing evidence further indicated that the receipt or improvement of ESG ratings is followed by a decline in idiosyncratic stock risk even after stringent controls for confounds—an effect that, in many studies, is materially driven by the social dimension (Horn, 2023).

Risk-tail evidence aligns with these mechanisms. Corporate social responsibility is associated with lower stock-price crash risk—i.e., fewer extreme downside realizations—suggesting that social performance curtails opacity-driven bad-news hoarding and reduces investors' sensitivity to adverse signals (Kim et al., 2014). The emerging-market evidence also documented a causal decline in idiosyncratic volatility as ESG performance improves, with information-environment upgrades cited as a principal channel; while composite, these effects are consistent with the idea that social practices reduce noise trading and stabilize expectations (Liu et al., 2023). Putting these strands together, theory and evidence jointly predict that higher social performance lowers firm-level volatility via human-capital reliability, stakeholder trust, disclosure-driven transparency, and tail-risk mitigation. Accordingly,

> **H2:** Social performance is expected to be associated with lower stock volatility, although the average linear effect may be limited in emerging-market settings.

ESG Performance, Stock Volatility and Governance Transparency

While prior hypotheses posit a negative association between ESG performance and stock volatility, the emerging evidence suggests that this relationship is not uniform across firms or market states. Instead, the ESG–risk nexus is nonlinear and state-dependent, reflecting heterogeneity in corporate governance, market sentiment, and institutional quality (Broadstock et al., 2019). Firms with stronger governance frameworks and transparent disclosure systems can credibly convey the informational content of their ESG activities, amplifying the risk-mitigating effects of sustainability practices. Conversely, when governance transparency is weak, ESG actions may be perceived as "cheap talk" or even as opportunistic greenwashing, attenuating or reversing the expected volatility reduction (Bae et al., 2017; Li et al., 2018).

From a theoretical perspective, signaling theory (Spence, 1973) provides a natural explanation for this heterogeneity. ESG engagement functions as a costly signal of firm quality only if external investors can verify its authenticity. In high-transparency settings, ESG signals reduce information asymmetry and uncertainty about the firm fundamentals, thereby lowering volatility. In low-transparency contexts, however, investors may discount ESG claims, generating a muted or even positive relationship with volatility. Agency theory further reinforces this conditionality. Strong governance aligns managerial incentives with long-term stakeholders, enhancing the credibility of sustainability disclosures and strengthening their stabilizing impact on market risk (Jensen & Meckling, 1976).

Empirical findings increasingly supported these nonlinear and conditional dynamics. Broadstock et al. (2019) found that the ESG–risk relationship intensifies during periods of high market stress, implying state-dependence driven by shifting the investors' attention to firm resilience. Bae et al. (2017) showed that in markets with better institutional quality, ESG performance has stronger value- and risk-implications, confirming governance transparency as a key moderating factor. Moreover, recent machine-learning evidence suggested that the marginal effect of ESG on volatility diminishes beyond a threshold, producing an Inverted-U shape: initial sustainability gains lower volatility, but beyond a point, excessive ESG investments may introduce uncertainty about capital allocation efficiency (Wang & Sonenshine, 2025).

In emerging markets, where governance quality varies widely, this nonlinearity becomes especially salient. Firms with robust disclosure and board independence enjoy lower ESG-related information risk, while those with opaque practices experience weaker or even reversed effects (Chao et al., 2022). Consequently, the risk-reduction benefits of ESG are state-contingent—amplified when governance transparency and institutional quality are high, and weakened when these are lacking. The prior literature suggested that the ESG–risk relationship may be nonlinear and state-dependent, particularly in environments where governance quality and disclosure credibility vary across firms. While governance transparency is theorized to condition the informational value of sustainability practices, data limitations prevent a direct empirical test of this moderating mechanism in the present study. Accordingly, nonlinearity and state-dependence were explored using learning-based models rather than explicit interaction terms.

Method

The study adopted an ex-post-facto, market-wide design to quantify how firm-level environmental and social practices relate to equity risk in an emerging market. The sampling frame comprised all non-financial companies listed on TSE between 2011 and 2023. Eligibility required (i) continuous listing throughout the window without prolonged trading suspensions or delisting, (ii) complete financial statements and market data necessary to compute returns, volatility, and controls, and (iii) auditable disclosure of sustainability practices in annual reports, notes, or stand-alone responsibility reports. Financial institutions (banks, insurers, investment companies, and holding firms) were excluded due to incomparable balance-sheet structures and regulatory regimes. Systematic screening yielded 109 firms that satisfied inclusion criteria; firm-year observations from these issuers constituted the unbalanced panel. The primary data sources were the CODAL disclosure system and Tehran Securities Exchange Technology Management Company for market microdata.

The firm-level environmental and social performance was operationalized via a transparent, replicable content-analysis checklist synthesized from the domestic literature and aligned with common ESG themes. Each disclosed practice received a binary score, with the environmental index aggregating items such as energy efficiency programs, greenhouse-gas reporting or reduction targets, waste management, compliance with environmental regulations, adoption of renewables, and green investment initiatives. The social index aggregated items such as workplace safety, employee training and development, workforce diversity and equal opportunity, customer-satisfaction programs, corporate philanthropy, and social transparency and accountability. While the checklist-based binary indices did not capture the intensity or qualitative depth of the sustainability actions, they offered transparency, replicability, and reduced measurement noise in settings where standardized ESG ratings are unavailable. In emerging markets characterized by heterogeneous disclosure quality, binary coding mitigates the subjective weighting and limits the researcher's discretion. Similar approaches were commonly adopted in disclosure-based ESG studies focusing on information availability rather than performance magnitude. For each firm-year, the environmental (E_Score) and social (S_Score) scores equaled the sum of disclosed items, producing bounded, interpretable indicators that are robust to scale differences across firms.

The market-based risks and control variables were constructed at matched frequencies. Annual stock volatility was defined as the standard deviation of firm-specific annual returns computed from closing prices; the market's conditional volatility was estimated at the firm level from daily returns using a GARCH (1, 1) specification, yielding an implied volatility proxy that captures time-varying risk. Additional controls followed the asset-pricing and corporate-finance literature: firm size (natural logarithm of total assets), financial leverage (total debt to total assets), broad market return (annual index return), and cash dividends per share. The firm-year stock return was retained both as a descriptive variable and, for reinforcement-learning reward design, as the realized payoff signal.

Preprocessing proceeded in four stages to ensure comparability and guard against artifacts. First, disclosures used in the checklist were dual-coded; disagreements were reconciled by adjudication to limit coder bias. Second, series were calendar-aligned to firm fiscal years; where necessary, price and dividend events were adjusted for corporate actions. Third, extreme observations in continuous financial variables were handled by moderate winsorization at conventional cutoffs after visual inspection via boxplots and distributional diagnostics. The ESG indices remained unaltered given their bounded nature. Fourth, continuous predictors were standardized within the training window only to avoid information leakage. The exploratory analysis assessed distributional shape (histograms, skewness, kurtosis), pairwise dependence (Pearson matrix), and missingness patterns; missing ESG items were treated as "not disclosed" by construction, while missing financials lead to firm-year omission. These steps, together with coding documentation, enabled full reproducibility.

Modeling followed a tiered strategy motivated by preliminary evidence that linear relations were weak while nonlinearity and temporal dependence were pronounced. As a baseline, multiple linear regressions related the stock volatility to E and S with controls, after verifying the standard assumptions. Residual normality was probed via Shapiro–Wilk tests and Q–Q plots, homoscedasticity via Breusch–Pagan tests and fitted-vs-residuals inspection, independence via the Durbin–Watson statistic, and multicollinearity via variance inflation factors; specification proceeded with heteroskedasticity-robust (HC) standard errors if needed. These diagnostics confirmed that low explanatory power is not driven by assumption violations and motivate non-linear learners.

The first machine-learning arm was a value-based reinforcement-learning framework that learned investment policies contingent on sustainability states. ESG information was discretized into five ordered states from very low to very high using quantile thresholds on the environmental score or the mean of E_Score and S_Score. The agent's action space was binary—invest versus skip—and the one-period reward equaled the realized firm-year stock return when investing and zero when skipping. Learning used tabular Q-Learning with ε-greedy exploration. Hyper-parameters were tuned in ranges customary for stable convergence in noisy financial environments: learning rate $\alpha \in [0.01, 0.10]$, discount factor $\gamma=0.90$, initial ε between 0.10 and 0.30 with geometric decay of 0.99, and 10,000 training episodes with randomized firm-year sampling. Convergence was monitored through the trajectory of cumulative reward and stabilization of Q-values across ESG states. The learned policy was extracted by greedy action selection on the terminal Q-table. The policy reasonableness was checked against economic priors and by sensitivity analyses that vary the ESG binning scheme and the reward definition to penalize volatility-seeking returns. Although the reward was defined as raw stock return, the reinforcement-learning framework was not intended to estimate risk-adjusted performance. Instead, it was used to examine whether ESG-related states contain systematic information that can guide sequential investment decisions. Risk implications were therefore inferred indirectly through state-dependent policy learning rather than explicit volatility penalization.

The second arm was a recurrent neural network that exploited the sequential nature of risk formation. A many-to-one LSTM architecture ingests multivariate firm-level sequences comprising lagged volatility, returns, and ESG scores along with time-varying controls, and outputs next-period volatility. Architectures with two to three stacked LSTM layers with 64–128 units each, tanh activations, and dropout between 0.2 and 0.5 fed into one or more dense layers. Weights were optimized with Adam at a 0.001 learning rate under a mean-squared-error objective. Training used mini-batches of 32–64 observations for 50–100 epochs under early stopping triggered by the validation RMSE. To preserve temporal causality, data were partitioned into contiguous train, validation, and test blocks with a rolling-origin scheme at the firm level; scalers were fitted on training windows only. The performance was summarized with MAE, RMSE, and out-of-sample R^2; learning curves were inspected for divergence to preempt overfitting. The model selection prioritized parsimonious architectures with the lowest validation RMSE and stable generalization.

The validation was deliberately layered. For the regression baseline, goodness-of-fit and inference relied on cross-sectional-time-series diagnostics and robust uncertainty estimates; for the reinforcement-learning policy, effectiveness was evaluated by comparing state-contingent Q-values and implied actions across ESG states and by tracking cumulative reward growth over episodes. For the LSTM, generalization was assessed via a strict out-of-sample test set never seen during training or hyper-parameter tuning; we additionally conducted a walk-forward experiment in which windows were advanced and models refitted to mimic real-time deployment. Across all learners, we computed bootstrap confidence intervals for key metrics and repeated the experiments under alternative random seeds to assess stability. Where appropriate, we contrasted the predictive accuracy against the linear baseline to quantify the incremental value of sequential and non-linear structure.

Robustness checks addressed construct validity and design choices. ESG states were re-defined using alternative cut points and an unsupervised k-means discretization; the reward was perturbed to risk-adjusted return (e.g., the return minus a volatility penalty) to gauge policy sensitivity; the LSTM input set was expanded with lagged GARCH volatility and pared back to ESG-only signals to isolate marginal contribution; horizons were shifted to test one- and multi-period forecasts; and leverage and size were re-scaled to mitigate the influence of outliers. The results were contrasted under firm and year fixed effects in the regression baseline to screen for latent heterogeneity. Collectively, these checks ensured that conclusions are not artifacts of particular measurement or modeling choices.

Implementation choices were justified by empirical diagnostics and the nature of the data. The binary ESG checklist yielded high inter-coder reliability and transparency in an environment where third-party ESG ratings were sparse. Weak linear associations observed in preliminary analyses supported the use of function approximators capable of learning non-linear and state-dependent relations. RL was appropriate when the estimand was a policy contingent on sustainability states, while LSTM networks were well-suited to capturing long-range dependencies and volatility clustering. The evaluation protocol emphasized out-of-sample performance and temporal integrity, aligning with best practice for financial prediction tasks.

Ethical and reproducibility safeguards were built in. All code implements deterministic data split with seeded randomness and logged configuration files for every run; preprocessing decisions, inclusion criteria, and checklist definitions were archived to enable audit. The analysis used only public disclosures and market data; no non-public or personally identifiable information was processed. Together, the sampling, measurement, modeling, and validation procedures constituted a coherent methodology for isolating whether and how firm-level environmental and social practices manifest in equity-market risk in an emerging-market setting.

Findings

Table 1 reports descriptive statistics for key variables including mean, median, standard deviation, minimum, maximum, skewness, and kurtosis. Skewness values clustered around zero indicate that the distributions were approximately symmetric. All kurtosis values were below 3, implying thinner-tailed distributions than the normal benchmark.

These distributional features suggest limited tail risk and provide a suitable basis for the subsequent modeling and inference.

Table 1.
The Descriptive Statistics of the Study Variables

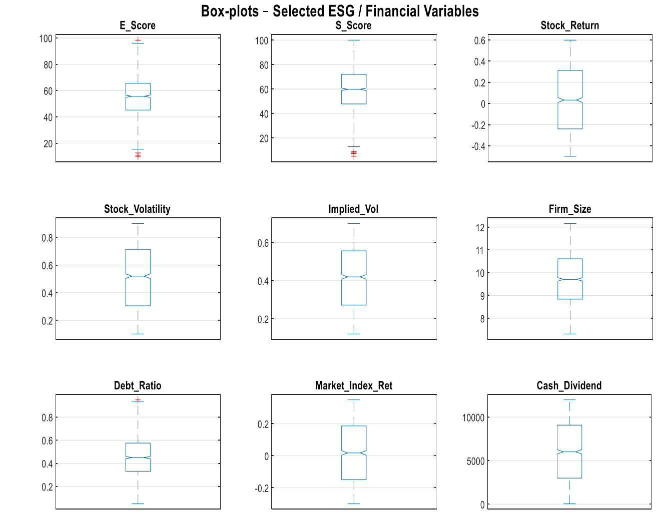
variable	Mean	Median	SD	Min	Max	Skewness	Kurtosis
E_Score	55.39	55.65	14.55	10.00	98.33	-0.07	2.84
S_Score	59.86	59.77	18.01	5.00	100.00	-0.03	2.79
Stock return	0.04	0.03	0.32	-0.49	0.59	0.07	1.80
Stock volatility	0.51	0.52	0.23	0.10	0.90	-0.03	1.76
Implied volatility	0.41	0.42	0.17	0.12	0.70	-0.06	1.82
Firm size	9.71	9.70	1.05	7.30	12.16	-0.00	1.94
Leverage	0.45	0.45	0.18	0.05	0.95	0.03	2.67
Market return	0.02	0.02	0.19	-0.30	0.35	0.05	1.80
Cash dividend	6068.73	6018.69	3443.21	48.07	11993.16	-0.00	1.81

Note: E_Score and S_Score are checklist-based indices scaled from 0 to 100.
(Source: The Researcher's Findings)

Table 1 presents descriptive statistics for the variables employed in this study, including environmental and social scores (E_Score and S_Score) as well as key market-based indicators such as stock return, stock volatility, implied volatility, firm size, leverage, market return, and cash dividend. The mean values of the environmental and social scores—55.39 and 59.86, respectively—suggest that the sampled firms, on average, maintained moderate-to-high levels of environmental and social engagement. The relatively high standard deviations (14.55 and 18.01) indicate substantial heterogeneity in ESG commitment across listed Iranian firms. From a market-behavior standpoint, the mean stock return of 0.04 with a relatively large standard deviation (0.32) implies positive but highly dispersed returns. The slight positive skewness (0.07) points to a mild right-tail bias, indicating that a limited number of firms experienced higher-than-average returns. The stock volatility and implied volatility had mean values of 0.51 and 0.41, respectively, both with a kurtosis below 2, reflecting concentrated distributions with limited extreme fluctuations. The firm size, with a mean logarithmic value of 9.71 and a kurtosis of 1.94, reveals that most firms in the sample were of medium to large scale. The average leverage ratio of 0.45 denotes a balanced capital structure, suggesting moderate reliance on debt financing. The market return, averaging 0.02, together with the mean cash dividend of 6,069 Toman per share, reflects the limited overall profitability of the TSE during the observation period. The analysis of skewness and kurtosis shows that the majority of variables followed approximately symmetric, thin-tailed distributions—conditions favorable for subsequent econometric and machine-learning modeling. These statistical properties ensure numerical stability and are particularly advantageous for reinforcement-learning and recurrent-neural-network training, which require well-behaved input distributions.

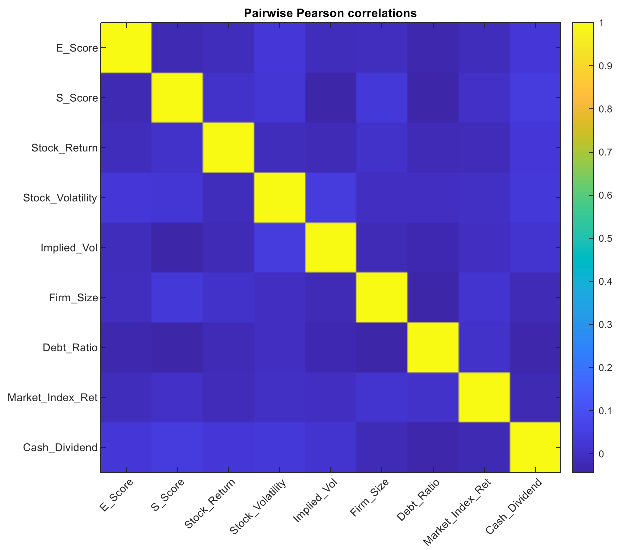
Figure 1 illustrates the distribution of the study's primary variables using box plots. Observable outliers appeared mainly in the E_Score and Cash_Dividend variables, reflecting the heterogeneity of environmental disclosure and payout policies across firms. For most variables, the median lay near the center of the box and the interquartile range was compact, suggesting approximately symmetric and near-normal distributions. These patterns indicate that the dataset was well-behaved and suitable for parametric modeling. The presence of moderate outliers provides valuable diagnostic insight for subsequent normalization and scaling procedures, thereby supporting the robustness of machine-learning algorithms—particularly reinforcement-learning and LSTM-based models—used in the later stages of analysis.

Figure 1.
The Box Plots of Key ESG and Financial Variables



(Source: The Researcher's Findings)

Figure 2.
Pearson Correlation Matrix of the Main Variables



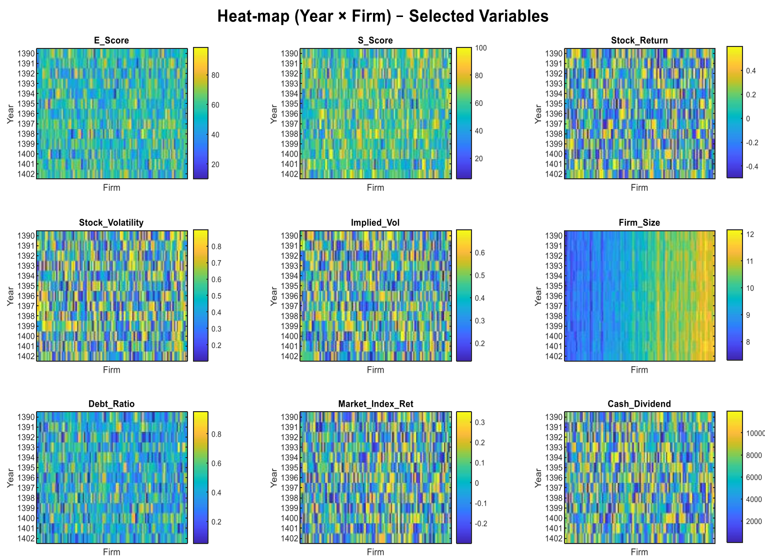
(Source: The Researcher's Findings)

Figure 2 presents the Pearson correlation matrix among the main variables of the study. The figure reports pairwise linear correlation coefficients and provides the same statistical information as a conventional correlation table, while offering a more intuitive visualization of the strength and direction of associations. As shown, the *E_Score* and

S_Score were positively correlated, reflecting internal consistency across sustainability dimensions. In contrast, ESG dimension did not exhibit a meaningful linear correlation with stock volatility or stock returns. The uniformly weak Pearson coefficients indicate that ESG–risk relationships did not operate through simple linear channels, thereby motivating the use of nonlinear and state-dependent modeling frameworks, such as RL and LSTM neural networks, in subsequent analyses.

Figure 3 visualizes the temporal and cross-sectional patterns of the study variables over the 2011–2023 period. Each panel depicts a two-dimensional heat map in which the color intensity represents the magnitude of a variable for each firm-year observation. The plots reveal substantial heterogeneity across firms and time, highlighting fluctuations in market behavior and sustainability performance. For instance, the Firm Size exhibited a relatively stable trajectory over time, reflecting the gradual evolution of corporate scale in the Iranian market, whereas the Market Index Return and Stock Volatility displayed more dispersed and dynamic patterns across firms. These variations capture both idiosyncratic shocks and sector-specific cycles that are essential for learning-based models to detect latent temporal dependencies. The heat maps provide an intuitive visualization of how ESG indicators and financial variables evolve jointly across time and firms, serving as an empirical foundation for reinforcement-learning and LSTM-based modeling in subsequent sections.

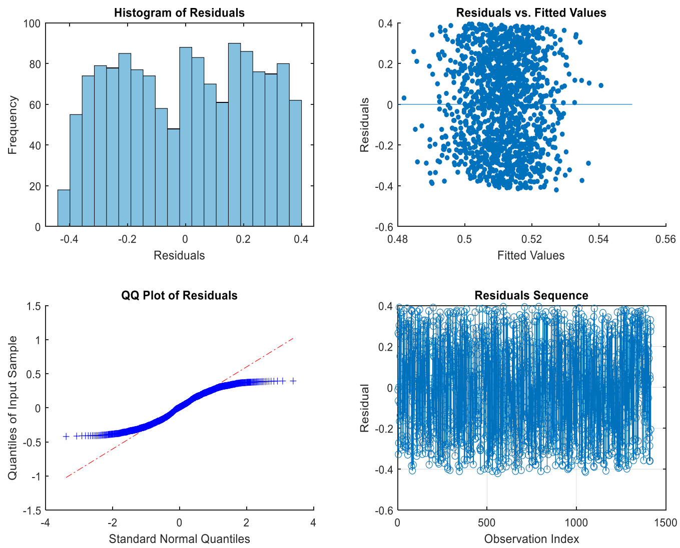
Figure 3.
The Heat Maps of Selected Variables by Year and Firm



(Source: The Researcher's Findings)

Figure 4 presents four diagnostic plots used to examine the assumptions of the classical regression framework. The upper-left panel shows a histogram of residuals, which approximates a normal distribution with a mild kurtosis. The upper-right panel (residuals versus fitted values) exhibited no discernible pattern, indicating homoscedasticity and the absence of systematic bias. The lower-left Q–Q plot demonstrated a slight deviation from the reference line at the tails, which may be attributed to a few outlying observations. The lower-right panel (residual sequence plot) revealed a random dispersion of residuals across observation indices, supporting the assumption of independence. Overall, the diagnostic results confirm that the regression model satisfied the key classical assumptions to an acceptable degree, validating its use as a benchmark for subsequent machine-learning analyses.

Figure 4.
The Diagnostic Plots of Residual Analysis for the Classical Regression Model



(Source: The Researcher's Findings)

Table 2 reports the Pearson correlation coefficients between the environmental and social scores (E_Score and S_Score) and stock volatility. The correlation coefficients were 0.02 for E_Score and 0.02 for S_Score, both positive but very close to zero. These results indicate the absence of a meaningful linear association between the firms' ESG performance and their stock return volatility. Furthermore, the p-values for both correlations exceeded the 0.05 significance level, suggesting that the observed relationships are statistically insignificant. Accordingly, it cannot be inferred that the ESG dimensions exerted a direct effect on market volatility. This finding aligns with several prior studies conducted in emerging markets, which suggest that the influence of ESG factors on firm risk is indirect or context-dependent. The weak linear correlations underscored the need to incorporate additional control variables and adopt more sophisticated modeling frameworks. Therefore, a multiple regression analysis was subsequently employed to examine the potential direct effects of ESG indicators after accounting for firm-specific characteristics.

Table 2.
The Pearson Correlation between ESG Scores and Stock Volatility

variable	Correlation Coefficient	p-Value
E_Score	0.02	0.30
S_Score	0.02	0.37

(Source: The Researcher's Findings)

Table 3 presents the results of the multiple regression model assessing the direct effects of ESG indicators (E_Score and S_Score) and control variables—firm size, leverage ratio, and market index return—on stock volatility. The findings reveal that none of the explanatory variables were statistically significant at the 95% confidence level. The coefficient of determination (R^2) was notably low (approximately 0.00), indicating that the model explained only a negligible portion of the variance in stock volatility. The overall model p-value exceeded the 0.05 threshold, confirming that the regression equation was not statistically significant. However, the Durbin–Watson statistic of 2.03 suggests that the residuals are free from serial correlation, implying no evidence of autocorrelation within the model. These results suggest that ESG indicators, in isolation and within this specific linear specification, did not meaningfully explain the fluctuations in stock volatility. The weak explanatory power may stem from unobserved latent factors, structural inefficiencies in the Iranian capital market, or potential time lags between corporate social responsibility actions and market reactions. Future research is encouraged to explore nonlinear or interaction-based modeling approaches—such as machine learning or reinforcement frameworks—to capture the complex dynamics between sustainability and financial risk.

Table 3.
The Results of Multiple Regression Explaining the Stock Volatility

variable	Coef.	95% CI (Lower)	95% CI (Upper)	VIF
Intercept	0.47	0.34	0.60	-
E_Score	0.00	-0.00	0.00	1.18
S_Score	0.00	-0.00	0.00	1.31
Firm Size	-0.00	-0.01	0.01	1.56
Leverage	-0.00	-0.07	0.06	1.45
Market Index Return	0.00	-0.06	0.06	1.05

(Source: The Researcher's Findings)

In addition to the regression analysis, the Variance Inflation Factor (VIF) was computed to examine the potential multicollinearity among the independent variables included in the model. All VIF values were found to be close to 1, and none exceeded the conventional thresholds of 5 or 10. These results confirm that no serious multicollinearity existed among the explanatory variables. Consequently, there was no concern regarding the coefficient instability or inflated standard errors due to internal correlations within the predictors. From a statistical perspective, the low VIF values strengthened the validity of the regression model by confirming the independence of the explanatory variables. The consistently low VIFs for the ESG indicators further suggest that the simultaneous inclusion of E_Score and S_Score did not introduce estimation bias or redundancy in the model. This diagnostic result is particularly important in multivariate analyses, where high inter-correlation can compromise the precision and interpretability of regression coefficients. The model structure demonstrates acceptable independence among the explanatory variables, ensuring reliable estimations for subsequent inferences.

The Reinforcement Learning Model: Q-Learning

The Q-Learning model is employed as a RL approach through which an autonomous agent interacts with the environment and iteratively learns optimal decisions over time. In this study, the environment was defined based on the firms' ESG performance. Specifically, the ESG scores were discretized into five ordered categories ranging from *very weak* to *very strong*. This state-space construction enabled the agent to evaluate and update its investment strategy across varying sustainability conditions. The rationale behind this design lies in the notion that firms with superior ESG profiles may exhibit more favorable long-term performance and risk characteristics. Thus, a state-based segmentation of ESG scores provides a structured foundation for learning ESG-driven investment policies.

The action set consisted of two choices: (0) do not invest, and (1) invest in the given firm. In the baseline specification, the reward was defined as the realized stock return: when the agent selected the investment action, the corresponding realized return was assigned as the reward; when the agent chose not to invest, the reward was set to zero. We note, however, that in markets with positive average returns this raw-return formulation can mechanically favor continuous investment and may therefore reflect a buy-and-hold tendency rather than any ESG-related risk mechanism.

To ensure that the reinforcement signal reflects the risk–return trade-offs, we adopted a risk-adjusted reward specification that penalizes conditional volatility:

$$Reward_t = r_t - \lambda.\, \sigma_t$$

Where r_t denotes the realized return in period t, σ_t is rolling (or conditional) volatility, and $\lambda \geq 0$ captures the degree of risk aversion. This modification allowed the agent to learn state-contingent policies that were informative about risk-sensitive investment decisions rather than raw performance alone.

By adopting this framework, the Q-Learning model facilitated the extraction of data-driven investment rules that reflected both financial performance and sustainability characteristics. As a result, the approach provides insights into whether responsible and ESG-oriented investment policies can be learned and systematically executed within a dynamic environment.

Table 4.
The Core Components of the Q-Learning Framework for ESG-Based Investment Analysis

Component	Definition	Implementation in This Study
State Space	Represents the underlying ESG condition of each firm, categorized into discrete sustainability levels (e.g., from very low to very high).	The firms' ESG scores are discretized into five ordered levels based on the environmental (E_Score) or the average of the environmental (E_Score) and social (S_Score), forming states 1 to 5.
Action	The decision made by the learning agent in each state, such as taking a position or remaining inactive.	Two discrete actions are defined: (0) do not invest and (1) invest in the firm.
Reward Function	Feedback received from the environment after taking an action, representing gains or losses associated with the chosen action.	The realized stock return (Stock_Return) serves as the reward. If the action = invest, the reward equals the firm's realized return; otherwise, the reward is set to zero.

(Source: The Researcher's Findings)

Table 5 reports the final Q-values obtained for the five discrete ESG states. A clear pattern emerges: across all states, the Q-value associated with the “Invest” action exceeded that of the “Skip” action. This indicates that, on average, taking an investment position generated higher expected cumulative rewards within the RL environment compared to remaining inactive. This result implies that, over the learning horizon, the agent identified investing in firms—regardless of their ESG tier—as a superior policy relative to staying out of the market. The highest Q-value for the invest action was observed in State 5 (the strongest ESG level), with a value exceeding 0.33, suggesting greater reinforcement and a stronger confidence signal toward investing in high-ESG firms.

Table 5.
The Final Q-Values for Discrete ESG States

ESG State	State1	State2	State3	State4	State5
Skip (0 = Do Not Invest)	0.01	0.01	0.01	0.01	0.01
Invest (1 = Invest)	0.21	0.18	0.27	0.23	0.33

(Source: The Researcher's Findings)

Table 6 presents the optimal policy learned by the Q-Learning algorithm across the five discrete ESG states. Consistent with the final Q-values, the model selected “Invest” as the optimal action in all ESG states. This indicates that, throughout the learning horizon, taking an investment position consistently yielded higher expected cumulative rewards relative to remaining out of the market. Importantly, this result reflects the agent’s policy under the specific reward structure and market data used in the study, where realized stock returns served as the reinforcement signal. While the policy implies that investing is systematically favored across all ESG levels, the stronger reinforcement observed at higher ESG tiers aligns with the notion that firms with robust sustainability characteristics may offer more favorable reward profiles. Nonetheless, these findings should be interpreted in the context of predictive policy learning rather than causal inference.

Table 6.
The Optimal Policy Learned Across ESG States

ESG State	State1	State2	State3	State4	State5
Optimal Action	Invest	Invest	Invest	Invest	Invest

(Source: The Researcher's Findings)

Table 7 reports the evolution of cumulative rewards across the training episodes in the Q-Learning model. At the beginning of the learning process, the cumulative reward was relatively low, approximately 0.85, reflecting the exploratory stage in which the agent has not yet learned an effective policy. As the training progresses, the cumulative reward steadily increased, reaching 10.15 by Episode 600. This upward trajectory indicates successful policy refinement and demonstrates that the agent gradually improves its investment decisions based on accumulated experience and feedback from the environment. The consistent increase in cumulative reward suggests that the algorithm effectively internalizes patterns from historical market behavior, allowing it to adopt more profitable strategies over time. While this result implies that ESG-based decision states provide useful signals within the RL framework, these findings should be interpreted as evidence of policy learning performance, rather than definitive proof of causal superiority of ESG-aligned investments.

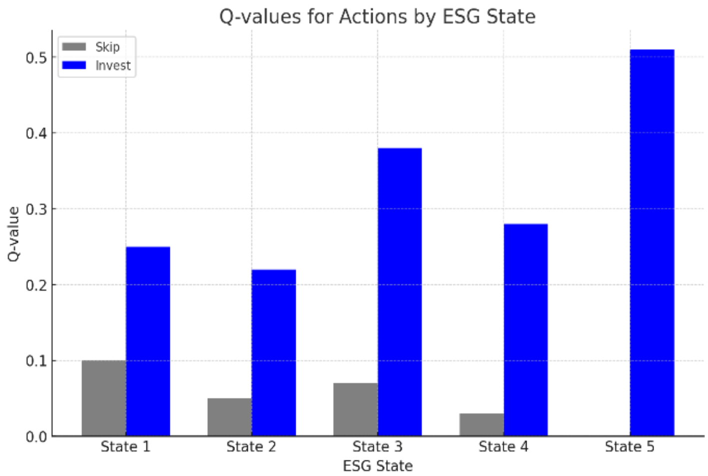
Table 7.
The Cumulative Reward Progression during the Q-Learning Training

Training Episode	1	100	200	300	400	500	600
Cumulative Reward	0.85	4.26	6.48	7.93	9.01	9.68	10.15

(Source: The Researcher's Findings)

This section summarizes the results of the Q-Learning model implemented to evaluate the investment decisions conditional on firms' ESG performance levels. The primary objective is to identify the optimal action—invest or abstain—across different sustainability states. To achieve this, Q-values were estimated for each action within each ESG tier and reported in tabular and graphical forms. Subsequently, the optimal policy was derived by selecting the action with the highest Q-value in each state. In addition, the learning dynamics were assessed by examining the trajectory of cumulative rewards over the training episodes, demonstrating the convergence of the agent's policy toward more profitable behaviors. These results illustrate that the RL framework can learn decision rules sensitive to ESG information and adaptively assign investment actions based on observed performance patterns. While these insights highlight the role of ESG signals in the learned policy, the findings should be interpreted as reflecting predictive investment behavior under the specified reward structure rather than establishing a causal superiority of ESG-compliant firms.

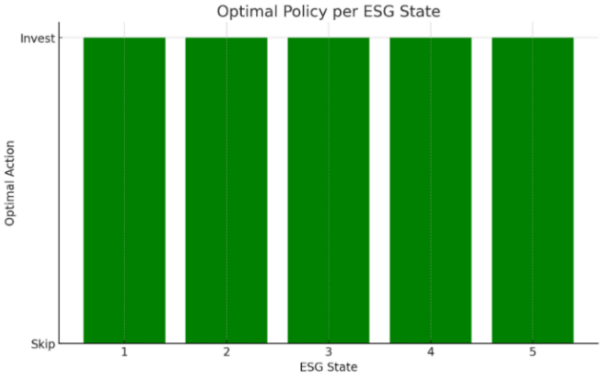
Figure 5.
The Comparison of Q-Values for "Invest" and "Skip" Actions Across ESG Levels



(Source: The Researcher's Findings)

Figure 5 presents the learned Q-values for the two possible actions—invest and skip—across the five ESG categories. In most states, the Q-value associated with the invest action exceeded that of the skip action, indicating that the agent, on average, expected higher cumulative returns when choosing to invest. This pattern suggests a reinforcement signal consistent with stronger ESG performance, as the agent increasingly favored investment decisions in higher-tier ESG states.

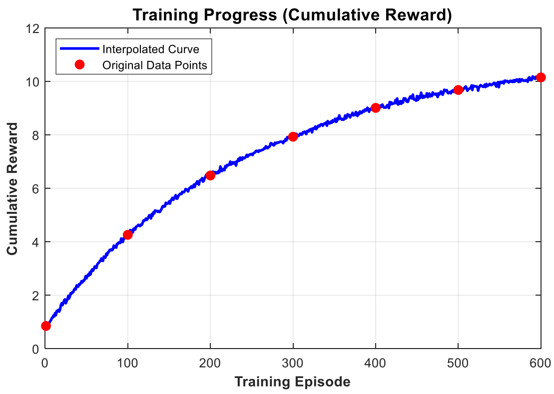
Figure 6.
The Optimal Policy across ESG States



(Source: The Researcher's Findings)

Figure 6 illustrates the optimal action selected by the Q-Learning model for each ESG state, where 0 denotes "skip" and 1 denotes "invest". The model consistently recommends investing across all ESG levels. This outcome reflects the learned policy under historical market dynamics and the adopted reward formulation, indicating that investing was reinforced as the dominant strategy throughout the training. Nonetheless, the result should be interpreted within the boundaries of the learning setup and does not imply universal dominance of ESG-driven portfolios in all market conditions.

Figure 7.
The Cumulative Reward Trajectory over the Training Episodes



(Source: The Researcher's Findings)

Figure 7 depicts the progression of cumulative rewards over the training episodes. The upward and stabilizing trend confirms that the model successfully converged toward an improved policy as training proceeded. The gradual increase in cumulative reward suggests the effective learning of investment behavior conditioned on ESG signals and demonstrates the model's ability to internalize performance feedback over time.

Recurrent Neural Network Modeling with LSTM

In this section, the performance of a Long Short-Term Memory (LSTM) neural network model is evaluated for forecasting the market volatility using ESG-based time-series inputs. The plotted results demonstrate a strong alignment between the actual volatility values and the model's predicted series, indicating that the LSTM architecture was capable of effectively capturing the temporal structure and dynamic patterns embedded in the data.

The evaluation table reports the model's predictive accuracy based on three commonly used metrics of Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2). It is important to emphasize that volatility is a highly persistent process in financial time series, and consequently, high predictive accuracy—particularly in terms of R^2—is commonly observed in models that effectively exploit lagged volatility information. Therefore, a high R^2 in this context should not be interpreted as evidence of superior structural modeling alone, but rather as a reflection of volatility clustering and temporal dependence. The reported values suggest that the model achieved a high degree of explanatory power and satisfactory predictive accuracy in reconstructing and forecasting stock-level volatility. In particular, the R^2 value approaching 1 reflected strong model fit under the specified input features and training configuration, although such performance should be interpreted within the controlled experimental setting and acknowledged as subject to typical time-series forecasting risks such as overfitting and regime variation. In this regard, the strong performance should be understood primarily as the model's capacity to capture the persistent volatility dynamics, rather than as an indication that all predictive power originates from the ESG-related inputs.

These results highlight the potential of ESG-enhanced sequential learning frameworks in modeling financial market volatility, supporting the notion that the sustainability-related signals may contain information relevant for forward-looking risk dynamics. However, this contribution should be viewed as incremental rather than dominant, as the bulk of forecasting power in volatility models typically arises from autoregressive dynamics and volatility clustering.

Table 8.
The LSTM Model Evaluation Metrics for Volatility Forecasting

Evaluation Metric	MAE	RMSE	R^2
Value	0.0583	0.0794	0.8835

(Source: The Researcher's Findings)

The results in Table 8 indicate the robust predictive performance of the LSTM model, with an R^2 of approximately 0.88, suggesting that the network explained around 88% of the variation in stock-level volatility under the given specification. The strong predictive performance of the LSTM model is partly attributable to the inclusion of lagged volatility, which is known to be highly persistent in financial time series. To further disentangle the contribution of ESG variables from that of lagged volatility and returns, an ablation-style comparison is conceptually informative. In particular, an LSTM specification excluding ESG inputs but retaining the same lag structure would be expected to preserve a substantial portion of predictive accuracy, given the volatility persistence. ESG variables are therefore interpreted as complementary signals that enhance volatility dynamics modeling rather than as primary drivers of predictive accuracy. The relatively low MAE and RMSE values further confirm that the model produced forecasts that are closely aligned with actual observations, reflecting effective learning and stable generalization across the dataset.

Table 9 compares the model performance across the training, validation, and testing phases. As expected, the model performed slightly better during the training; however, the performance gap remained small across all three data segments. This consistency indicates satisfactory generalizability and suggests that the model is not subject to substantial overfitting.

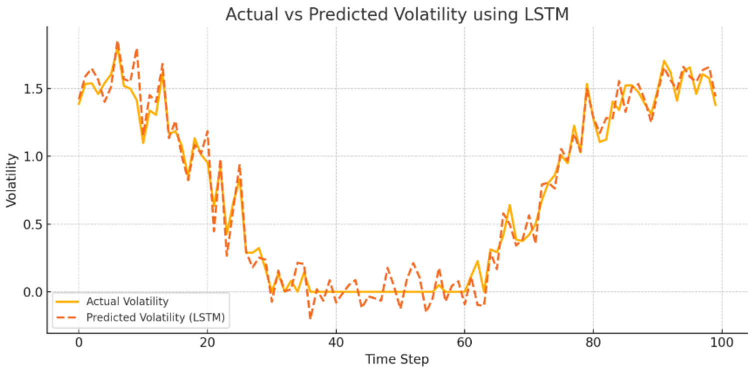
Table 9.
The Comparison of Performance across Training, Validation, and Test Sets

Data Split	MAE	RMSE	R^2
Training	0.05	0.07	0.89
Validation	0.05	0.07	0.86
Test	0.06	0.08	0.84

(Source: The Researcher's Findings)

Figure 8 compares the actual stock volatility series (solid orange line) with the LSTM-predicted values (dashed light-orange line) over a rolling window of 100 time steps. The close co-movement between the two curves—particularly during pronounced upward and downward movements (steps 0–30 and 60–100)—demonstrated the model's ability to capture nonlinear market dynamics. Minor deviations at sharp turning points largely reflected sudden market shocks not fully represented in the training distribution. Overall, this visual evidence further supports the LSTM model's reliability in forecasting volatility for ESG-oriented equities.

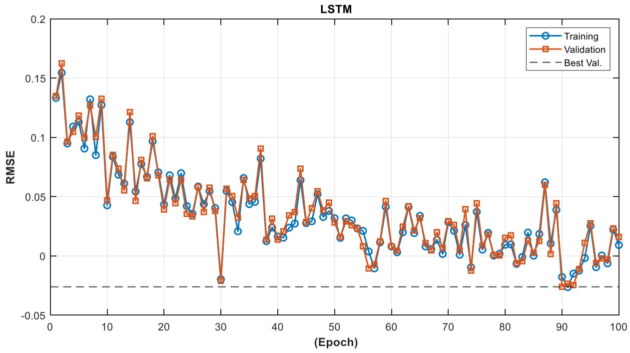
Figure 8.
The Training and Validation RMSE across Epochs



(Source: The Researcher's Findings)

Figure 9 illustrates the RMSE trajectory over 100 training epochs for both the training (blue line) and validation (orange line) datasets. The steady decline in error values until approximately epoch 60 reflected a progressive improvement in the model's predictive capability. Beyond this point, both curves fluctuated slightly while remaining closely aligned, indicating a stable convergence and the absence of severe overfitting. The dashed black vertical line marked the minimum validation RMSE, serving as a reference point for potential early stopping and optimal weight selection. These results suggest that the chosen architecture and learning rate configuration are appropriate for the ESG-driven volatility forecasting task.

Figure 9.
The Learning Curve of the LSTM Network



(Source: The Researcher's Findings)

To prevent the information leakage, data splitting was conducted along the time dimension within firms. For each firm, observations were partitioned into non-overlapping training, validation, and test windows in chronological order. No future information enters the model estimation or hyperparameter tuning, ensuring a strictly out-of-sample evaluation.

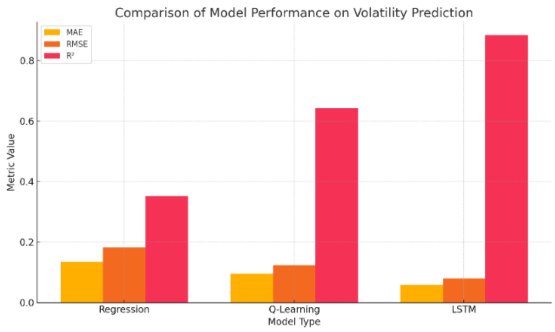
As shown in Table 10 and Figure 10, the LSTM model demonstrates superior performance across all evaluation metrics relative to the benchmark models. As shown in Table 10 and Figure 10, the LSTM model demonstrates strong predictive performance relative to the classical regression benchmark. It is important to clarify, however, that Q-learning was not a volatility forecasting model but a policy-learning framework. As such, its performance metrics were not directly comparable to forecasting accuracy measures derived from regression or LSTM models. The comparison was therefore intended to highlight methodological differences and complementary analytical roles rather than to establish a strict ranking of predictive models.

Table 10.
The Comparative Performance of Alternative Models in Volatility Forecasting

Model	MAE	RMSE	R^2
Classical Regression	0.08	0.10	0.26
Q-Learning (RL)	0.06	0.09	0.64
LSTM Neural Network	0.05	0.07	0.88

(Source: The Researcher's Findings)

Figure 10.
The Comparative Performance of the Models



(Source: The Researcher's Findings)

Based on the comparative results presented, it is evident that machine-learning and reinforcement-learning approaches—particularly the LSTM network—demonstrate superior performance relative to the traditional statistical model. Based on the comparative results presented, machine-learning approaches—particularly the LSTM network—exhibited strong predictive performance in volatility forecasting tasks. This performance primarily reflected the ability of deep sequential architectures to exploit temporal dependence and volatility persistence. ESG-related variables provided incremental informational content but should not be interpreted as the primary source of predictive accuracy. RL, in turn, served a complementary role by illustrating state-dependent decision mechanisms rather than by offering a direct forecasting benchmark.

Discussion and Conclusion

The empirical evidence obtained in this study indicated that the environmental and social performance do not exhibit statistically significant average linear associations with firm-level stock volatility in an emerging market context. The near-zero correlations and statistically insignificant coefficients observed in traditional regression models indicated that ESG variables, when modeled linearly, provide limited explanatory power for stock volatility; however, the learning-based models revealed that sustainability-related information can be informative for capturing nonlinear and dynamic volatility patterns beyond the average linear effects. These findings align with stakeholder theory (Donaldson & Preston, 1995; Freeman, 1984) and agency theory (Jensen & Meckling, 1976), both of which posited that enhanced transparency and alignment of managerial decisions with stakeholder interests bolster investor confidence and thereby reduce market uncertainty.

While the linear models did not provide evidence of statistically significant average effects, the learning-based results suggested that environmental information may be more informative than social information in capturing nonlinear and dynamic volatility patterns. This difference should be interpreted in a predictive and exploratory sense rather than as evidence of a causal or statistically significant risk-reduction effect. The environmental practices may be more tightly linked to operational processes and regulatory exposure, which learning-based models can exploit when modeling volatility dynamics, whereas social practices may influence market perceptions more gradually and indirectly.

The validation of the non-linearity and state-dependence hypothesis constitutes one of the most impactful contributions of this research. The learning-based analyses pointed to nonlinear and state-dependent patterns in the ESG–volatility relationship. While prior literature suggested that governance transparency may condition the informational value of sustainability practices, this study did not directly test the governance-based moderation due to data limitations. Accordingly, references to governance were interpreted conceptually, serving to contextualize the observed nonlinear dynamics rather than to establish a tested conditional or risk-mitigating mechanism. This pattern dovetails with findings from Rahat and Nguyen (2024) and Biju

(2025), which highlighted that the informational value of ESG disclosures is contingent on institutional credibility.

The integration of traditional econometric models with artificial-intelligence techniques allowed us to capture both linear relationships and temporal, non-linear dependencies—addressing a lacuna in the extant ESG-finance literature. The superior predictive performance of the LSTM model relative to linear regression underscored the importance of modeling the dynamic, non-linear nature of risk in sustainability-driven environments. The implications are clear. Policymakers and regulators should prioritize the development of standardized ESG-reporting frameworks to enhance transparency and reduce information asymmetry. Investors may incorporate ESG metrics into risk assessment frameworks and view them not only as ethical indicators but also as predictive indicators of volatility. Corporate managers should conceptualize ESG engagement as a strategic instrument—rather than a mere reputational tool—for stabilizing performance and attracting long-term capital.

This study offered one of the first comprehensive examinations of how firm-level environmental and social performance influence stock volatility in an emerging-market context via a hybrid econometric–machine-learning approach. The findings indicated that the sustainability practices do not exert statistically significant average linear effects on equity risk; instead, their relevance emerges through nonlinear and time-dependent patterns captured by learning-based models. Environmental performance appeared more informative than social performance in learning-based models of volatility dynamics; however, neither dimension exhibited statistically significant average linear effects on equity risk, and the overall the ESG–risk nexus exhibited nonlinear and state-dependent patterns, which were explored through learning-based models rather than through explicit governance-based moderation tests. Theoretically, the findings are consistent with stakeholder, agency, and signaling perspectives by suggesting that sustainability-related disclosures may contribute to transparency and legitimacy, which learning-based models can exploit when modeling nonlinear and time-dependent volatility dynamics, without implying statistically significant average linear risk-reduction effects. Methodologically, this work advances the ESG literature by utilizing RL and LSTM architectures, which outperform conventional linear models in capturing temporal and non-linear features of financial volatility. However, this study is not without limitations. The reliance on manually coded ESG disclosures and the focus on a single emerging market limit generalizability. Future research may extend the present framework in several important directions. One promising avenue is to conduct industry-specific analyses in order to explore potential sectoral heterogeneity in the relationship between ESG performance and stock volatility, particularly in settings where environmental and social risks are unevenly distributed across industries. Such extensions would allow researchers to examine whether the nonlinear and state-dependent effects documented in this study vary systematically across sectors with different regulatory exposure, competitive dynamics, or sustainability pressures. In addition, future studies may broaden the empirical scope to cross-country settings, explicitly incorporate governance dimensions as a separate analytical component, and employ standardized ESG ratings where data availability permits. Further research could also investigate asymmetric effects of sustainability performance during market downturns versus expansionary phases, as well as pursue causal identification through quasi-experimental or natural-experiment research designs.

REFERENCE

Albuquerque, R., Koskinen, Y., & Zhang, C. (2019). Corporate social responsibility and firm risk: Theory and empirical evidence. *Management Science, 65*(10), 4451–4469. https://doi.org/10.1287/mnsc.2018.3043.

Bae, K.-H., El Ghoul, S., Guedhami, O., Kwok, C., & Zheng, Y. (2017). Does corporate social responsibility reduce the costs of high leverage? Evidence from capital structure and product market interactions. *Journal of Banking and Finance, 100*, 135–150. https://doi.org/10.1016/j.jbankfin.2018.11.007.

Biju, A. V. N. (2025). ESG-firm performance nexus: Evidence from an emerging economy. *Business Strategy and the Environment*. *34*(3), 3469-3496. https://doi.org/10.1002/bse.4152.

Broadstock, D. C., Matousek, R., Meyer, M. T., & Tzeremes, N.G. (2019). Does corporate social responsibility impact firms' innovation capacity? The indirect link between environmental & social governance implementation and innovation performance. *Journal of Business Research, 119*(2), 99-110. https://doi.org/10.1016/j.jbusres.2019.07.014.

Chao, L., Mian, W., Xi, C., & Wenli, H. (2022). Environmental, social and governance performance, corporate transparency, and credit rating: Some evidence from Chinese A-share listed companies. *Pacific-Basin Finance Journal*, *74*(C), 101806. https://doi.org/10.1016/j.pacfin.2022.101806.

Chava, S. (2014). Environmental externalities and cost of capital. *Management Science, 60*(9), 2223–2247. https://doi.org/10.1287/mnsc.2013.1863.

Dhaliwal, D. S., Li, O. Z., Tsang, A., & Yang, Y. G. (2011). Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate social responsibility reporting. *The Accounting Review, 86*(1), 59–100. https://doi.org/10.2308/accr.00000005.

Donaldson, T., & Preston, L. E. (1995). The stakeholder theory of the corporation: Concepts, evidence, and implications. *Academy of Management Review, 20*(1), 65–91. https://doi.org/10.5465/amr.1995.9503271992.

Edmans, A. (2011). Does the stock market fully value intangibles? Employee satisfaction and equity prices. *Journal of Financial Economics, 101*(3), 621–640. https://doi.org/10.1016/j.jfineco.2011.03.021.

Farūq, A. T. M., & Chowdhury, M. A. R. (2025). Financial markets and ESG: How big data is transforming sustainable investing in developing countries. *arXiv*. https://arxiv.org/abs/2503.06696.

Freeman, R. E. (1984). *Strategic management: A stakeholder approach*. Pitman.

Gupta, H., & Chaudhary, R. (2023). An analysis of volatility and risk-adjusted returns of ESG indices in developed and emerging economies. *Risks*, *11*(10), 182. https://doi.org/10.3390/risks11100182.

Horn, M. (2023). The influence of ESG ratings on idiosyncratic stock risk: The unrated, the good, the bad, and the sinners. *Schmalenbach J Bus Res, 75*, 415-442.

https://doi.org/10.1007/s41471-023-00155-1.

Iannone, B., Duttilo, P., & Gattone, S. A. (2025). Evaluating the resilience of ESG investments in European markets during turmoil periods. *arXiv*. https://arxiv.org/abs/2501.03269.

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics, 3*(4), 305-360. https://doi.org/10.1016/0304-405X(76)90026-X.

Kansoy, F., & Stasiulaitis, D. (2025). Green shields: The role of ESG in uncertain time. *arXiv*. https://arxiv.org/abs/2506.02143.

Kim, Y., Li, H., & Li, S. (2014). Corporate social responsibility and stock price crash risk. *Journal of Banking & Finance, 43*, 1–13. https://doi.org/10.1016/j.jbankfin.2014.02.013.

Li, Y., Gong, Z., Zhang, X., & Koh, L. (2018). The impact of environmental, social, and governance disclosure on firm value: The role of CEO power. *British Accounting Review, 50*(1), 60–75. https://doi.org/10.1016/j.bar.2017.09.007.

Lins, K. V., Servaes, H., & Tamayo, A. (2017). Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. *The Journal of Finance, 72*(4), 1785–1824. https://doi.org/10.1111/jofi.12505.

Liu, D., Gu, K., & Hu, W. (2023). ESG performance and stock idiosyncratic volatility. *Finance Research Letters, 58*(B), 104393. https://doi.org/10.1016/j.frl.2023.104393.

Naseer, M. M., Guo, Y., Bagh, T., & Zhu, X. (2024). Sustainable investments in volatile times: Nexus of climate change risk, ESG practices, and market outcomes. *International Review of Financial Analysis*, *95*(B), 103492. https://doi.org/10.1016/j.irfa.2024.103492.

Rahat, B., & Nguyen, P. (2024). The impact of ESG profile on firm's valuation in emerging markets. *International Review of Financial Analysis*, *95*(A), 103361. https://doi.org/10.1016/j.irfa.2024.103361.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics, 87*(3), 355-374. https://doi.org/10.2307/1882010.

Valencia Söderberg, D., & Truong, M. (2024). *ESG & emerging markets – A volatility perspective of ESG investment in emerging markets.* Linköping University.

Wang, H., Shen, H., & Li, S. (2023). ESG performance and stock price fragility. *Finance Research Letters, 56*, 104101. https://doi.org/10.1016/j.frl.2023.104101.

Wang, Y., & Sonenshine, R. (2025). The nonlinear impact of environmental, social, governance on stock market performance among US manufacturing and banking firms. *Journal of Risk and Financial Management*, *18*(6), 293. https://doi.org/10.3390/jrfm18060293.

Journal of
Knowledge Economy Studies

https://kes.hmu.ac.ir

Hazrat-e Masoumeh University

A Survey on a Conceptual Model of Enterprise

Article Type:
Research Article

Zeinab Rajabi[1*]
Department of Computer Engeenering, Hazrat-e Masoumeh University, Qom, Iran
Corresponding Author
E-mail: z.rajabi@hmu.ac.ir

Seyed Mohsen Rahnamafard[2]
University of Tehran,Tehran, Iran
E-mail: rahnamafard@alumni.ut.ac.ir

Received: 11 September 2025
Revised: 25 October 2025
Accepted: 31 December 2025
Available Online: 01 January 2026

Abstract

Enterprise ontology serves as a foundational framework for semantically comprehending the nature of organizations and the essential components that uphold their integrity. The systematic and conceptual understanding of organizations has garnered significant attention from researchers due to its pivotal role in various domains, including business modeling, enterprise architecture, business process management, context-aware systems, application development, interoperability across diverse systems and platforms, knowledge management, organizational learning and innovation, and conflict resolution within organizations. Achieving a consensus on the concepts related to the fundamental elements that constitute an organization is therefore critical. This study aimed to conduct a comprehensive analysis and comparison of the existing conceptual models of enterprises as documented in scholarly articles published over the past decade. The comparison revealed significant variations in coverage, adaptability, and maturity across models, with many lacking completeness or alignment with comprehensive frameworks like Zachman's framework. The strengths and weaknesses of each model were discussed and a robust framework for their evaluation was introduced. To facilitate this evaluation, we proposed several pertinent criteria derived from established methodologies for assessing the ontologies. Furthermore, we identified contemporary challenges and issues that have been overlooked in prior studies, offering insights and suggestions for future research directions in enterprise modeling. Ultimately, a roadmap for enhancing the systematic understanding of organizations through refined enterprise ontology frameworks was presented.

Keywords

Enterprise ontology, Conceptual model, Enterprise dimension, Enterprise architecture.

Cite this article: Rajabi, Z. & Rahnamafard, M., (2026). A survey on a conceptual model of enterprise. *Journal of Knowledge Economy Studies (JKES)*, 3(1), 73-93. DOI: http//doi.org/10.22034/kes.2026.2071314.1080

Online ISSN: 3060-7329

1. Introduction

Scholars have provided numerous definitions for the term organization in literature. Most defined an organization as a group of individuals who have come together to achieve a specific goal (Daft, 2020). An organization is a phenomenon whose most important constituent is human beings. According to postmodernist beliefs shared by several philosophers over the past couple of decades, enhancing interactions among humans is one of the most important areas for achieving high-performance organizations (Stacey et a., 2000). One relevant topic in this area is semiotics, which concerns interactions formed through signage and sign processes, including modeling languages (Mingers, 2006).

On the other hand, according to systems theory, an organization is also one of the many systems around us. Most scholars whose ideas are based on Newtonianism and systematic thinking believed that an organization is a system consisting of elements between which distinct relationships exist; thus, understanding organizations depends on grasping these elements and their relationships (Robbins, 1990). Despite these scientific approaches, there is still a lack of a common language that would help bring about a general consensus in the research community (Senge, 1990).

Management scholars have developed models to better understand organizations, their organizational problems, and to help classify and interpret their data systematically (Rock & Crawford, date?). Some studies (e.g., Burke & Litwin, 1992; Nadler & Tushman, 1995; Porras & Robertson, 1986; Waterman Jr et al., 1980; Weisbord, 1976) modeled the relationships between organizational components. For example, Levitt (1965), in his diagnostic organizational model, introduced organizational components such as structure, technology, people, and tasks. Weisbord presented a six-box organizational cognitive model (Saleem & Ghani, 2013; Weisbord, 1976) which includes purposes, structure, relationships, rewards, leadership, and helpful mechanisms. Janicijevic (2010) examined and compared the organizational cognitive model, dividing organizational elements into two categories of static and dynamic. Static elements include organizational structure, systems, culture, informal groups, and power structure, whereas dynamic elements are business processes, group processes, leadership, conflicts, political processes, and communication. Having reviewed the literature, he concluded that the existing diagnostic organizational models are imperfect because they do not include dynamic formal organizational components like business processes. According to him, a complete and comprehensive diagnostic model should encompass business processes and related parameters such as process owners and participants, organizational competence, indicators of key performance, shortcomings and problems of business process, key paths to change business processes, and priorities of business process.

On the other hand, information technology researchers want to align information technology with the goals of the organization and also develop information systems after the organization's structured recognition. Artificial Intelligence (AI) researchers also seek to understand the structure of organizations with the aim of creating a suitable environment for their own systems. Ontology is the most relevant area of science that aims to recognize these aspects of organizations. This field was first introduced by AI experts in order to make sense of human semantic treasure for machines. The field of ontology has developed methods and tools to build different conceptual models for verbal and non-verbal concepts in various subject domains, one of which includes that of organizations. Rosing (2015) reviewed business ontology research and examined how business ontology is used in organizational development. He used the potential of ontology and semantics to develop standards that describe objects, relationships, and rules for enterprise modeling, organizational engineering, and enterprise architecture.

Enterprise ontology provides a uniform representation of similar semantic content (Dietz, 2006). Modellers use different methods to develop models. These models are created with different languages and modeling tools. There may even be various styles and different techniques used within a single method. In addition to this, products created by different organizations and disciplines use different terms to analyze organizations, leading to various perceptions of the organization. Therefore, a standardized format is needed to translate data among different systems of the organization and to understand different models of organizational analysis (zur Muehlen, 2009).

Enterprise ontology provides a data structure that facilitates the reader's understanding of data usage in an organization description document. For example, Rajabi et al. (2013) presented the methodology for enterprise architecture development based on enterprise ontology. The ontology of the enterprise provides the necessary information to collect, organize, and store data in an easy way to understand (Kindrick, 2009; Rajabi & Abade, 2012). For example, the Dodaf Data Meta Model (Model) indicates that the goal of a conceptual model is to support the integrity and semantic accuracy of architectural descriptions.

On the other hand, the enterprise ontology helps to model more efficiently by describing the building blocks of enterprise and their relationships. The enterprise ontology is a proper basis for an integrated understanding of an organization's elements. The enterprise ontology actually models the building blocks of organizations with their relationships according to the perception of entities from two parties (zur Muehlen, 2009). The relationships among all elements of the organization are modeled precisely, transparently**,** and are formulated in the ontology of the organization. Then, a common model which has the necessary precision for all parties within the organization and systems is created.

The advantages and successful applications of ontology in business and various applications are quite clear for researchers (Feilmayr & Wöß, 2016). The ontology development for organizations is the proper basis for enterprise architecture methods (Hinkelmann et al., 2016; Rajabi et al., 2013; Rajabi & Abade, 2012),() automatic analysis of models at enterprise architecture, querying and inference in architectural data (Antunes et al., 2014), business process management (Jung, 2009; Rao et al., 2012; Santos Jr et al., 2010),) business modeling (Gassen et al., 2017), business process re-engineering (AbdEllatif et al., 2018; Rao et al., 2012),) implementation of applications (Villela et al., 2005), context-aware systems (Aguilar et al., 2018), interoperability between different systems and platforms (Chen et al., 2008), and knowledge management in the organization (Villela et al., 2005). Therefore, it is of paramount importance to identify a suitable ontology that has the necessary comprehensiveness, proper coverage, accuracy, compatibility, and extensibility for several applications.

This study aimed to evaluate and compare enterprise ontology models from the conceptual view and then analyze their results. O'Leary (2010) reviewed the enterprise ontology according to activity theory but he does not consider many other aspects of the enterprise ontology. Besides O'Leary's work, it could be said that there are no other proper comparisons and classifications of enterprise ontology models available. Therefore, researchers who need to use the enterprise ontology models in different domains may become confused as the domain of relevance of each model is not clear.

In this paper, a conceptual framework is proposed to compare the enterprise ontologies. Relevant evaluation criteria are identified and applied to analyze the existing ontology models. Based on these analyses, key insights are derived and a roadmap for future research on conceptual models of enterprise ontology is presented.

2. Related Works

2.1 Ontology

According to Gruber's definition in 1993 (Gruber, 1995), an ontology is a formal, explicit specification of shared conceptualization. On the basis of this definition, "conceptualization" refers to an abstract model of phenomena in the world along with the detection of related concepts to those phenomena. "Explicit" means that the types of used concepts and their limitations are defined explicitly. "Formal" refers to the fact that the ontology should be readable to a machine and "shared" indicates that the ontology must acquire agreed and acceptable knowledge by related societies (Antunes et al., 2014). Although this definition emphasizes the formal and explicit description of concepts, these descriptions need to first agree on selected concepts and an acceptable conceptual model. If the concepts are not chosen appropriately, the ontology usage will not be efficient. A well-defined conceptual model is useful in many research studies and applications independently.

2.2 Conceptualization

A formal model is implemented in an ontology language such that the ontologist observes a gradual transition from the knowledge level to the implementation level. The formalization grade of the knowledge model increases gradually until it is able to be understood by the machine. Figure 1 shows this gradual movement.

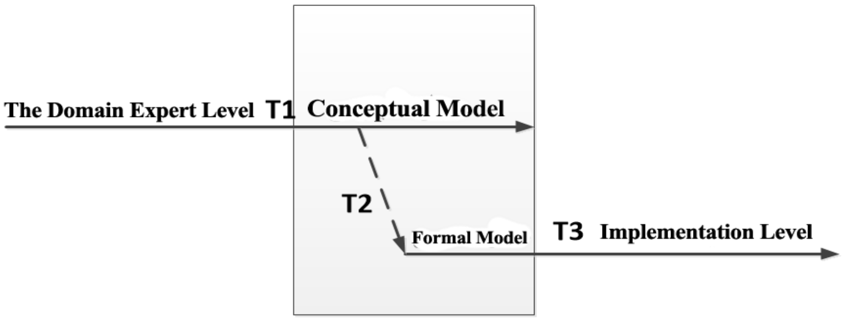
The ontology development activities generally include: specification, conceptualization, formalization, implementation, and maintenance. Conceptualization is a crucial activity in the ontology development process (Gómez-Pérez, 1996). Some studies emphasized this and provided some methods for conceptualization. The conceptualization activity constructs meaningful knowledge models from the domain knowledge. The conceptualization activity is similar to collecting puzzle pieces provided by the knowledge acquisition activity, and it is completed during the conceptualization process. The conceptualization activity must be done rigorously; otherwise, the error will propagate into the next steps.

The purpose of conceptualization is to prepare a domain model with a lower degree of formality than a formal model but still more formal than the definition of the model in natural language. Other motivations for the conceptualization process include:

1. Domain experts, human users, and ontologists may struggle to interpret or understand the ontology proposed in the ontology language.
2. Domain experts may not be able to construct ontologies within their domain of expertise.

This activity deserves special attention because it plays an important role in the ontology development process. In ontology development methodologies, after the conceptual model has been designed, the conceptual model is transformed into a formal model and implemented in an ontology language such that the ontologist observes a gradual transition from the knowledge level to the implementation level. The formalization grade of the knowledge model increases gradually until it is able to be understood by the machine. Figure 1 shows this gradual movement.

Figure 1.
Knowledge in the Ontology Development Process



(Gomez-Perez et al., 2006)

T1 transformation refers to the conceptual modeling process that transforms the domain expert's subjective model into a conceptual model. T2 refers to the conceptual model progressing into a formal model. T3 refers to the formal model progressing into a model that can be understood by a machine. As the figure shows, T1 and T3 transformations are drawn by a continuous line, while the T2 transformation is indicated by a dashed line. This indicates that there may be some loss of domain knowledge when a conceptual model is developed into a formal model. This happens when the components and tools used to create conceptual models are more meaningful and expressive than those that are used to create the formal model.

In the methodology development process (Fernández-López et al., 1997), the conceptualization activity uses a set of intermediate representations (IRs) through table and graph notations, organizing and converting an informal representation of a domain into specifications that can be understood by domain experts and ontology developers. In this study, we aimed to compare the enterprise ontology models from a conceptual perspective.

A conceptual model is designed for human understanding and explanation of domain knowledge, often represented using diagrams like ER or UML. These models are simple, intuitive, and suitable for analysis and communication among stakeholders. A formal ontology, in contrast, provides a precise, machine-processable representation using logical languages such as OWL or RDF. Its purpose is to enable automated reasoning, consistency checking, and knowledge sharing. The conceptual model usually serves as a preliminary step, providing a foundation for building a formal ontology.

2.3 Ontology-Based Enterprise Modeling

The enterprise ontology contains a set of well-defined terms that are widely used as common descriptions of enterprises, as it accurately covers concepts related to the enterprise field. The enterprise ontology acts as an interactive medium or platform between different people, such as users, designers, and planners in various organizations (Uschold et al., 1998).

An important issue in achieving integration and performing effective business planning is that all operators and stakeholders (from planning managers to low-ranking contractors involved in software production) must have a common understanding of different dimensions of an organization. When a particular word is used from the domain, the concept it refers to should be clear. In other words, it is necessary to overcome the "semantic heterogeneity" associated with implicit perceptions of common words' meanings in the domain.

The enterprise ontology has been created for this purpose and contains a set of well-defined words that are widely used as general descriptions of an organization, covering concepts related to the domain of the organization carefully. This set facilitates a shared understanding of an organization and can serve as a fixed basis for identifying the functional requirements and creating the organizational models. Thus, perceptual errors are reduced in cases where the same concepts may be referred to by different terms, as it improves and facilitates the interaction between stakeholders, which is an important step in increasing efficiency.

The purpose of applying an ontology in an organization is to determine the relationships between the tasks and activities of that organization with organizational knowledge and their tools. It also participates in the acquisition, representation, and manipulation of organizational knowledge, organized and structured libraries of the existing knowledge, rationalized descriptions of inputs and outputs of involved components, and a vocabulary exchange format for an enterprise (Ciocoiu et al. , 2001).

3. Investigating the Conceptual Models of the Enterprise Ontology

Researchers represented different ontology models for different applications. This article selects the pioneering research such as TOVE[1] ((Fox et al., 1995; Gomez-Perez et al., 2006), context-based enterprise ontology (Leppänen, 2007), and the enterprise ontology TEO[2] (Uschold et al., 1998). These projects and other researchers are recognized as pioneers in this article. In addition, the enterprise architecture frameworks require a conceptual model of an organization. Therefore, some

1. Toronto Virtual Enterprise
2. Time Event Ontology

frameworks and methodologies provided conceptual models of enterprise ontology for developing enterprise architecture, such as the Dodaf[1] Data Meta Model (Officer, 2009; Thakor & Sasi, 2015),"DoDAF PLUGIN user guide" (2010), Modaf[2] (Aue & Gamon, 2005), and the Togaf[3] Content Meta Model (Pereira & Almeida, 2014). ArchiMate (Pereira & Almeida, 2014) was also investigated in this study. Some researchers focused only on one dimension of an organization and presented an enterprise ontology model for it, as referred to in Table 1. For example, Almeida and Gizzareti, Pereira and Almeida (2014), Santos et al. (2013), Abramowicz et al. (2008), and Pereira (2015) have modeled the organizational structure ontology and represented extensive details. Some other studies constructed an enterprise ontology model for a specific type of organization. In this regard, Silva and Belo (2018) have presented an enterprise ontology model specifically for higher education institutions.

Table 1.
The Enterprise Ontology Which has Examined the Concepts of One of the Dimensions of the Organization

Dimension	References
Structure	(OMG, 2009), (Santos Jr et al., 2013), (Almeida & Guizzardi, 2008), (Pereira & Almeida, 2014), (Abramowicz et al., 2008), (Diorbert Corrêa Pereira, 2015), (Carvalho & Almeida, 2015)
Purpose	Business Motivation Model (BMM) (OMG, 2015)
Rules and Time	Date-Time Foundation Vocabulary Request For Proposal (OMG, 2008)

(Source: The Researcher's Findings)

The best examples of such research can be found in the documentation provided by the Object Management Group (OMG). Several OMG documents serve as valuable references for analyzing the conceptual model of enterprise ontology, presenting concepts in an organized manner and detailing their relationships with one another. One notable document is the Business Motivation Model (BMM) (OMG, 2015), which offers a structured set of concepts that model elements of a business plan. This document is particularly useful for identifying the purpose and motivation of an organization, including concepts such as purpose, mission, perspective, and strategy. Additionally, OMG has published a comprehensive document on business process modeling known as BPMN[4]. This document not only covers conceptual control flow modeling but also provides precise definitions of components such as activities, events, gateways, and the sequences between them. It also explores the relationships between various organizational components and activities. For example, it represents capital resources using Pools and Lanes while defining consumable resources with a symbol called the Data Object. Furthermore, the Organization Structure Meta-model (OSM) (OMG, 2009) document from OMG offers metamodeling for organizational structures. It includes modeling elements that represent organizational entities, subgroups, features, and the relationships between organizational units and their assigned individuals. The concepts used in the organizational structure are clearly defined within the OSM document. The Semantics of Business Vocabulary and Rules (SBVR) (Kang et al., 2010) takes this a step

1. Department of Defence Architecture Framework
2. Ministry of Defence Architecture Framework
3. The Open Group Architecture Framework
4. [Https: //www. omg. org/spec/BPMN/2. 0/PDF]

further by expressing business vocabularies with formal logic, providing a specific language for business descriptions. It defines a set of terms, each with a specific technical meaning relevant to the field of business. The rules defined by formal logic closely resemble natural language, making SBVR a business ontology that machines can understand (Kang et al., 2010). Lastly, the Date Time Foundation Vocabulary Request for Proposal (OMG, 2008) is another OMG document that articulates concepts related to time and dates in business using SBVR.

3.1 Providing a Framework for Comparison

There has been no existing frameworks to compare the conceptual models of enterprise ontology until now, making our study the first to address this issue. In this regard, we presented a framework for evaluating the conceptual models of enterprises. Two perspectives have been considered in formulating this framework, which includes various comparison parameters. The first perspective aims to identify the closest semantic frameworks to enterprise ontology and draws inspiration from their parameters for comparison. The second perspective identifies and applies general parameters for evaluating ontologies, supported by detailed research in the field. In first perspective, the most significant research was conducted by Osterwalder (Gordijn et al., 2005), who presented a framework for comparing business model ontologies based on earlier works (e.g., Jasper & Uschold, 1999; Pateli, 2003). Our study generalized the Osterwalder's framework to facilitate the comparison of enterprise ontologies. This generalization seeks to provide an acceptable framework for comparing ontologies that model organizations. In this framework, we introduced important parameters for the comparison of enterprise ontology, which are described as follows:

1. **Purpose:** This parameter reflects the primary motivation behind the enterprise ontology and the objectives of its development. Several studies (e.g., Officer, 2009), presented ontology models specifically for applications in enterprise architecture. Others (e.g., Poels et al., 2018), concentrated on ontology models tailored for business contexts. Additionally, some studies (e.g., Leppänen, 2007) offered a more general enterprise ontology aimed at enterprise modeling.
2. **Domain:** This parameter evaluates the domain of relevance for the conceptual model. The enterprise ontology model is capable of representing various types of enterprises, including business, military, and educational organizations.
3. **Implementation language:** This parameter indicates the programming languages used to implement the enterprise ontology and convert it into a machine-readable format. It encompasses the use of generic ontological technologies for representing ontologies, such as Ontolingua, RDF/S, and OWL, as well as ontology design tools like Protégé.
4. **Representation:** Representation distinctions form a core aspect of ontological classification. Lightweight ontologies focus on establishing a basic semantic structure by defining concepts, their taxonomies (e.g., Subclass-Of relationships), the relationships between them (e.g., part-of), and the properties (slots) that describe each concept. They are primarily concerned with vocabulary and a simple

hierarchy. In contrast, heavyweight ontologies extend this foundation by incorporating formal axioms and constraints. These logical rules explicitly define the intended meaning of the vocabulary and restrict its interpretation, enabling complex consistency checking and automated reasoning.

This distinction is best illustrated with an example. Consider an ontology for a corporate domain. A lightweight version might define concepts like Employee, Manager, and Department, with relationships such as worksIn (Employee, Department) and subClassOf (Manager, Employee). This is often represented using software engineering diagrams like UML Class Diagrams or Entity-Relationship Diagrams (ERD). The corresponding heavyweight ontology would add axioms like:

$\forall x$ Manager(x) $\rightarrow \exists y$ manages(x, y) (Every Manager manages at least one other Employee).

Disjointness constraints: Manager and Department are disjoint (no individual can be both).

A cardinality constraint: An Employee works in exactly one Department.

These axioms, typically implemented using AI-based formalisms like Description Logics (the foundation for OWL - Web Ontology Language), eliminate ambiguities and allow a reasoner to infer new knowledge like detecting an inconsistency if an individual is asserted to be both a Manager and a Department.

5. **Ontology content and components:** This parameter refers to the key dimensions addressed by each enterprise ontology. Each conceptual model of an organization encompasses various dimensions and the concepts associated with them. It is crucial to identify which dimensions are included at macro level and which concepts are most critical among all the concepts. Additionally, this parameter takes into account the types of relationships and the nature of the rules presented within the ontology.
6. **Ontology maturity and evaluation:** The degree of maturity of an ontology is determined by its evaluation. Various qualitative criteria and resources exist for assessing the ontological models that have been implemented (e.g., Brank et al., 2005; Hlomani & Stacey, 2014; McDaniel et al., 2018). Several criteria possess significant capability for studying the evaluation of conceptual models. In this article, we selected specific criteria to understand and compare the existing ontological models of the organizations. If the evaluation criteria function effectively at the conceptual level, we can anticipate that the enterprise ontology will perform well at the formal level.

   It should be noted that the relative importance of the evaluation criteria introduced in this study—such as extensibility, consistency, and completeness—is not the same across all applications. In fact, the weight and priority of each criterion depend on the purpose and domain in which the ontology is applied. For example, in architecture enterprise, completeness and consistency are of greater importance, whereas in applications such as context-aware systems or business process management, extensibility and adaptability may be more critical. Therefore, the proposed framework does not assume equal weighting of the criteria; rather, it allows for contextual adjustment of their relative importance according to the specific application domain.

The selected criteria for evaluation are as follows:

a) **Reusability:** Reusability refers to the extent to which the components of an ontology can be applied in developing another ontology for a different domain or purpose. This criterion assesses how modular and generalizable the ontology is. For example, if a "process" or "actor" concept from an organizational ontology can be directly reused in designing a project management ontology without major modifications, it demonstrates high reusability. Since ontology development is often complex and time-consuming, the ability to reuse the existing structures significantly reduces effort and promotes consistency across the related domains.

b) **Accuracy:** Accuracy refers to the degree to which the knowledge represented in the ontology reflects the real-world understanding of domain experts. An ontology should faithfully capture the actual entities, relationships, and constraints that exist in the domain it represents. For example, in an organizational ontology, the concept of *"employee"* must accurately include attributes such as *position, department,* and *reporting relationships*, consistent with how these elements are defined and used in the real organization. Even though ontologies inherently allow for interpretive flexibility, maintaining close alignment with expert knowledge ensures that the model remains valid and useful for real-world applications.

c) **Expandability:** Expandability refers to the ability of an ontology to be extended or enriched with new concepts and relationships without altering its existing definitions or structure. A highly expandable ontology provides a flexible foundation that can evolve as new requirements or related domains emerge. For example, an organizational ontology that defines general concepts such as *"role", "process",* and *"resource"* can later be expanded to cover domains like *human resource management* or *knowledge management* by adding new subclasses or relationships—without the need to modify the original definitions.

d) **Adaptability:** Adaptability refers to the extent to which an ontology can accommodate or anticipate future changes in its domain or environment. It evaluates whether the ontology offers a stable yet flexible foundation that can evolve as new requirements, technologies, or organizational structures emerge. For example, an organizational ontology that models communication channels in a generic way—like defining a general concept of "interaction" instead of specifying only "email" or "meeting"—can easily adapt to future developments like AI-driven collaboration tools or virtual workspaces without major restructuring. Many existing ontologies lack this level of flexibility, making them difficult to maintain or extend over time.

e) **Completeness:** Completeness measures the extent to which an ontology provides full coverage of its target domain, ensuring that it contains all essential concepts, relationships, and properties needed to answer relevant questions within that domain. An ontology is considered complete when it can represent the domain comprehensively without major conceptual gaps. For example, an organizational ontology would demonstrate completeness if it includes all key elements such as *employees, departments, goals, processes,* and *resources*, enabling it to support queries like "which department is responsible for a specific goal?" or "what resources are involved in a given process?". The parameters are summarized in Table 2. This table, along with its parameters, provides a framework for comparing the enterprise ontologies, which we will discuss in this section.

Table 2.
The Ontology Evaluation Parameters That are More Aligned with Conceptual Evaluation Model

Parameter		References	Description
Purpose		(Gordijn et al., 2005)	The main motivation of organizational ontology and the purpose of creating the ontology.
Domain		(Gordijn et al., 2005)	The domain of organization which the ontology is modeling, for example business, military or educational organization.
Implementation language		(Gordijn et al., 2005)	The implementation language and applied language used to create the enterprise ontology.
Representation		(Gordijn et al., 2005)	How the the enterprise ontology model is represented.
Ontology content and component		(Gordijn et al., 2005)	The dimensions of the domain considered by the conceptual model and the concepts underpinning them.
Ontology maturity and evaluation	Reusability	(McDaniel et al., 2018)	The degree to which the entire ontology or part thereof can be repurposed and reconstructed another ontology.
	Accuracy	(Burke & Litwin, 1992)	The degree of consistentency of the ontology with the knowledge of a domain expert.
	Expandability	(Gómez-Pérez, 1996)	The ability to extend the ontology to other domains without changing definitions.
	Adaptability	(McDaniel et al., 2018)	Whether the model reacts predictably towards the small internal changes or not.
	Completeness	(Burke & Litwin, 1992)	The ability to how exhaustively the ontology can answer all questions that ontology should be able to answer.

(Source: The Researcher's Findings)

3.2 A Comparison of Enterprise Ontologies

In this section, we compare the conceptual models according to our framework presented in the previous section.

Purpose: The main motivation of organizational ontology and the purpose of creating the ontology. A number of studies such as TOGAF Content model (Awadallah, 2013; Weisman, 2011)), ArchiMate (Pereira & Almeida, 2014; Wierda, 2017), DODAF Data Meta Model (Officer, 2009), UAF[1](OMG, 2017a; OMG, 2017b) provided ontology models for applications of the enterprise architecture. Some studies such as context-based (Leppänen, 2007), The Enterprise Ontology (TEO) (Uschold et al., 1998), and TOVE (Fox & Gruninger, 1998) presented the enterprise ontology in general for enterprise modeling.

Domain: The domain refers to the specific type of organization that the ontology is designed to model. This could encompass various sectors such as business, military, or educational organizations. Each domain has its unique characteristics, structures, and processes, which the ontology aims to represent accurately. By tailoring the ontology to a particular domain, it becomes more relevant and useful for stakeholders within that field, facilitating better understanding, communication, and decision-making.

1. Unified Architecture Framework

The TOGAF Content Model (Awadallah, 2013; Weisman, 2011), ArchiMate (Wierda, 2017), and the Unified Architecture Framework (UAF) (OMG, 2017a; OMG, 2017b) are generic frameworks in the realm of enterprise ontology. In contrast, other models (Poels et al., 2018) focus specifically on ontology models for the business domain. Additionally, the DODAF Data Meta Model (Officer, 2009) provides an ontology model tailored for the military domain.

Implementation language: Most ontology models such as Dodaf Data Meta Model (Officer, 2009), TOGAF Content Model (Awadallah, 2013; Weisman, 2011) , ArchiMate (Wierda, 2017), UAF (OMG, 2017a; OMG, 2017b) are represented at the conceptual level by UML diagrams. TOVE (Fox & Gruninger, 1998) is implemented by Prolog language and TEO (Uschold et al., 1998) is implemented by Ontolingua language (based on KIF).

Most ontology models, including the DODAF Data Meta Model, TOGAF Content Model, ArchiMate, and UAF, are typically represented at the conceptual level using UML diagrams. In contrast, TOVE is implemented using the Prolog programming language, while TEO is developed using Ontolingua which is based on KIF.

Representation: Most ontology models are implemented in a lightweight form, while only TOVE and TEO (Uschold et al., 1998) are classified as heavyweight ontology models.

Content and componentsb The core conceptual model identifies the main dimensions of an organization, followed by the detailed concepts that support each dimension. For example, TEO (Uschold et al., 1998) represents five key dimensions of activity, organization, strategy, marketing, and time.

Reusability: Most enterprise ontologies immediately transition into the implementation phase without first establishing a solid conceptual model. This oversight limits the users' ability to connect the abstract concepts of the model with the real-world elements they are intended to represent. A robust conceptual model is crucial for supporting reusability. Among organizational ontology models, both TOVE and TEO rush into implementation, leaving users without a comprehensive understanding of the models, which hampers effective usage. In contrast, the DODAF Data Meta Model offers a well-defined conceptual model that articulates relationships at a conceptual level, although it is specifically tailored for the United States Department of Defense. The Context-Based Enterprise Ontology starts with a conceptual level presentation, but many of its relationships remain unclear, limiting its reusability. Meanwhile, the UAF describes each concept precisely, supporting extensibility; however, it suffers from ambiguities at macro level, making reusability challenging.

Accuracy: TOVE and TEO exhibit ambiguous concepts, with their definitions and relationships not being clearly defined. This leads to varying interpretations of each concept. In contrast, the DODAF Data Meta Model provides a well-defined enterprise ontology aimed at representing the conceptual model of defense organizations. Additionally, the UAF, which originated from DODAF and MODAF, is designed to support non-defense organizations. The TOGAF Content Meta Model clearly defines concepts and their relationships, offering a solid foundation for enterprise ontology criteria; however, it does not implement the ontology model at a formal level.

Expandability: Context-based enterprise ontology, TOGAF Content Meta Model, ArchiMate, and UAF are generally defined in a way that allows for good expandability into specific domains. In contrast, the relationship between "activity" and "capability" in the DODAF Data Meta Model is tailored specifically for military organizations, which limits its applicability to other sectors despite its well-defined concepts.

Adaptability: TOVE and TEO transition abruptly into the formal phase without adequately defining their concepts. On the other hand, the DODAF Data Meta Model excels in defining the conceptual phase but is primarily suited for military organizations. The TOGAF Meta-Model considers strong concepts at micro level, yet it lacks a comprehensive ontological structure.

Completeness: Completeness refers to how thoroughly an ontology can address all relevant questions about the organization it represents. This means that the ontology should cover all dimensions of an organization. Given the social nature of organizations, researchers must consider multiple dimensions simultaneously. For instance, defining “service” requires acknowledging the roles of both service customers and providers. Similarly, a complete understanding of a “business process” necessitates describing the roles of its participants. To achieve this, the structure of organizational units and the roles within them must be framed within a broader organizational context. However, focusing on too many dimensions can lead to selecting concepts that may not be essential for a complete description of the domain. An ontology can be considered comprehensive if it effectively answers questions such as: who (performer) does what (task) for what reasons (goal), where (location), and when (time) (Leppänen, 2007)?

Table 3.
The Comparison of Enterprise Ontology According to the Proposed Framework

		TOVE	**The Enterprise Ontology(TEO)**	**Context-based**	**DODAF Data Meta Model**	**TOGAF Content Model**	**ArchiMate**	**UAF**
		(Fox & Gruninger, 1998)	**(Uschold et al., 1998)**	**(Leppänen, 2007)**	**(Officer, 2009)**	**(Awadallah, 2013); (Weisman, 2011)**	**(Pereira & Almeida, 2014); (Wierda, 2017)**	**(OMG, 2017b); (OMG, 2017a)**
Purpose		Enterprise modeling	Enterprise modeling	Enterprise modeling	Enterprise Architecture	Enterprise Architecture	Enterprise Architecture	Enterprise Architecture
Domain		Public and commercial	Commercial Enterprise	Public	Military	Public	Public	Public
Implementation language		Prolog	Ontolingua (Base on KIF)	At the conceptual level and UML	At the conceptual level and UML	At the conceptual level and UML	has provided its own modeling language	At the conceptual level and UML
Representation		Heavyweight	Heavyweight	lightweight	lightweight	lightweight	lightweight	lightweight
Ontology content and components		Organization, Resource, Activity, Time, Cost	Organization's Activity, Strategy, Marketing, Time	Purpose area, Actor area, Action area, Object area, Facility area, LOcation area, Time area	Activity, Capability, Resource (Information, Performer and Material), Location, Guide	Governance, Service, Process, Data, Infrastructure, Motivation	3 layers of business, application and technology that stand under each concepts	Taxonomy Structure, Connectivity Processes, States Interaction Scenarios, Information Constraints, Roadmap Traceability
Ontology maturity and evaluation	Reusablity	Mid	Mid	High	Mid	Mid	Mid	Mid
	Accuracy	Low	Low	Mid	Mid	High	High	Mid
	Expandability	Low	Low	High	High	High	High	High
	Adaptability	Low	Low	Low	Mid	High	High	High
	Completeness	Low	Low	Mid	High	High	High	High

(Source: The Researcher's Findings)

3.3 Investigating the Ontology of Enterprise Completeness Using the Zachman Framework

Studies on enterprise ontology examined various dimensions of organizations. For instance, Leppanen introduced a context-based ontology (Leppänen, 2007) that encompasses seven dimensions of goal, actor, action, object, facility, location, and time. The TOVE project (Fox et al., 1995; Fox & Gruninger, 1998) considered four dimensions of ontology of organization (Fox & Gruninger, 1998), ontology of resource (Fadel et al., 1994), ontology of activity (Gruninger & Fox, 1994), and ontology of cost (Tham et al., 1994). ArchiMate (Wierda, 2017) presented a meta-model structured into three layers of business, application, and technology. Additionally, ArchiMate categorized its elements into three groups of active elements that perform actions, behavioral elements that represent the behavior of active elements, and passive elements that are acted upon by behavioral elements. The TOGAF Content Model (Awadallah, 2013; Weisman, 2011) outlined concepts such as motivation, infrastructure, data, process, service, and governance.

The Zachman Framework (Zachman, 1987; Sowa & Zachman, 1992) aimed to examine all dimensions of an organization, making it a suitable foundation for studying the models of enterprise ontology. The columns of Zachman's Framework represented various aspects (dimensions) of an organization, derived from the 5W1H questions of who (responsibilities), when (time), why (motivation), where (location), how (task), and what (data). A central question arises: does the ontology cover all dimensions of the enterprise? If it does, then the enterprise ontology will demonstrate good comprehensiveness. The six communication questions of 5W1H help clarify the dimensions of organizations, as noted by Caetano et al. (Caetano et al., 2012) and Zachman. Rajabi et al. (2013) presented an enterprise ontology model based on the factors in the columns of Zachman's Framework. By utilizing the 5W1H questions, Zachman clarified each dimension of the organization, providing a solid basis for understanding the existing ontologies and their covered dimensions. In this paper, we compared the completeness of the enterprise ontology models using the columns of Zachman's Framework.

In addition to reviewing the concepts of enterprise ontology (see Table 3), we also asseseds their compatibility with Zachman's Framework. The selected ontologies for comparison are leading models that have sufficient documentation available. For instance, the TOGAF Content Meta-Model introduces important concepts such as data entity, value stream, constraint, role, organization unit, location, business service, process, function, and business capability. Each of these concepts is comprehensively defined along with their relationships to one another. Within the Zachman's Framework, the concepts can be categorized as follows: data entity and value stream fall under the "what" column; process, function, and business service are placed in "how" column; location is categorized under "where"; and organization unit is found in "who" column. Notably, there are no concepts represented under the "when" column. As illustrated in Table 3, the Context-Based Enterprise Ontology defines specific areas for each column of Zachman's Framework and outlines various concepts for each area. This model demonstrates better alignment with the columns of Zachman's Framework, enhancing its adaptability and relevance.

Some ontology models include "business product" as a key concept within the enterprise ontology. For instance, ArchiMate (Wierda, 2017) considered "product" as an essential element of the organization, defining it as anything offered to the outside world. This definition also encompasses products that may be provided internally to different parts of the organization. Thus, the concept of "product" is crucial, yet it is overlooked in some other ontologies.

The concept of "location" is included in the ontology of organizations, but some models, such as DODAF, TOGAF, and ArchiMate, limit their definition to a single concept of "location". In contrast, the Context-Based Enterprise Ontology and the UAF consider multiple concepts related to location. On the other hand, both TOVE and enterprise ontology do not address any concepts for the "where" column. The concept of "business service" in TOGAF (Awadallah, 2013) supports business capabilities through an explicitly defined interface and is governed by an organization. Similarly, the UAF defines "service specification" as a set of functionalities provided by one element to be used by others. This indicates that the concept of "service" is significant within the enterprise architecture; however, it is absent from models that focus solely on organizational modeling. Table 4 summarizes these findings regarding the adaptability within Zachman's Framework, while Table 5 presents the final results for the adaptability of the existing models according to Zachman's Framework.

Table 4.
A Review of the Adaptability of the Existing Concepts in Enterprise Ontology Within Zachman's Framework

Enterprise ontology	Why	What	How	Who	Where	When
TOVE (Fox & Gruninger, 1998)		Goal, Sub goal	Activity, Constraint, Authority Communication link	Resource, Organization, Division, Subdivision, Team, Agent, Role, Skill		
The Enterprise ontology (Uschold et al., 1998)	Purpose, Mission	CSF, Objective, Vision, Goal	Activity, Activity Spec, Sub-Activity, Execute, Plan, Sub-Plan, Process Spec, Org. Structure, Strategy, Risk, Capability	Entity, Role, Relation Attribute Resource, Person, Corporation, Unit, Actor, Machine, Actor Role, Skill, Activity Owner, Doer, Authority		Time Point, Time Interval, T-Begin, T-End, Time Line, Calendar, Date, Duration
Context-based Enterprise Ontology (Leppänen, 2007)	Reason, Purpose	Goal	Function, Activity, Task, Action Structure	Facility, Resource, Tool, Human actor, Person, Group, Position, Role, Unit, Organization	Location Area, Physical Location Point, Spatial Thing Logical, Location Region, Geographical Dimension, Geographical System	Time, Time Point, Time Interval, Time Unit, Time System, Clock Time, Calendar Time
DODAF Data Meta Model (Officer, 2009)		Vision, Desired Effect	Project, Capability, Activity, Guidance, Condition	Personnel Type, Skill, Performer, Data, Information, Materiel	Location	
TOGAF Content Meta Model (Awadallah, 2013); (Weisman, 2011)		Business Service	Function, Process, Value Stream, Course of Action, Business Capability, Constraint	Organization Unit, Function, Role, Data Entity	Location	
ArchiMate (Pereira & Almeida, 2014); (Wierda, 2017)		Business Service, Business Product	Business Process, Business Function, Business Interaction, Business Event, Contract	Business Role, Business Actor, Business Collaboration, Business Object	Location, Business Interface	
UAF (OMG, 2017b); (OMG, 2017a)		Service, Service Specification, Enterprise Vision, Enterprise Goal	Capability, Project Kind, Project Activity, Project Milestone, Capable Element Project, Actual Milestone Kind, Operational Activity	Actual Organization, Organizational Resource, Person, Post, Responsibility, Natural Resource, Physical Resource, Resource Architecture, Resource Artifact, Resource Performer, Software System, Standard, Protocol, Protocol Stack	Location, Location Holder, Location Kind, Actual Location	

(Source: The Researcher's Findings)

Table 5.
The Results of the Adaptability of the Existing Concepts in Enterprise Ontology Within Zachman's Framework

Conceptual Model	TOVE	The Enterprise Ontology (TEO)	Context-based	DODAF Data Meta Model	TOGAF Content model	ArchiMate	UAF
Adaptability with Zachman's framework	×	×	✓	×	×	×	×

(Source: The Researcher's Findings)

4. Findings

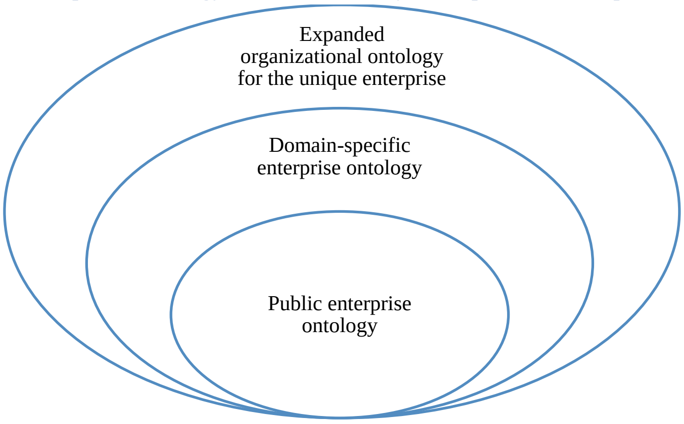
Numerous studies have explored the enterprise ontology; however, many exhibited significant weaknesses. A critical issue is the lack of consensus on which dimensions of an organization should be included in its ontological model. For instance, while one study (e.g., Fox & Gruninger, 1998) incorporated the time dimension, another study (e.g., Officer, 2009) overlooked it entirely. Establishing a common agreement on the conceptual model is essential before progressing to formal and logical construction phases. We require a foundational conceptual model that accurately represents key concepts and relationships, yet currently, there are no standard or reference models available for researchers to consult.

The existing enterprise ontology models often lack conceptual depth, leading to premature transitions into the implementation phase. This results in underdeveloped conceptual models, making their formal counterparts neither reusable nor expandable. Furthermore, comprehending the formal models becomes challenging without a robust conceptual framework.

Additionally, the concepts present in current ontologies do not adequately encompass all the components of an organization. If ontological models were better aligned with the columns outlined in Zachman's framework, they could more effectively cover various organizational dimensions. There is an urgent need for a powerful conceptual model that articulates fundamental concepts and relationships in an interpretable manner for diverse applications, including enterprise architecture, business architecture, business process management, context-aware systems, intercommunication, and automated production and analysis of the models. Moreover, there is no standardized approach to expand and customize a generic enterprise ontology model for specific domains or organizational needs.

The conceptual models of enterprises can be enriched progressively. Initially, a general enterprise ontology model is developed. Then, it is refined for specific industries. Ultimately, an enterprise-specific ontology emerges based on the previous models. This progression is illustrated in Figure 2. In general, all organizations share a set of common principles and concepts that form their foundation. In the second stage, these common concepts are detailed to suit various types of organizations—such as commercial entities, military organizations, and universities. Finally, in the third step, appropriate common concepts are tailored specifically to describe individual organizations.

Figure 2.
The Relationship Between Enterprise Ontology and Its Subsets (Development and Expansion of Ontology)



(Source: The Researcher's Findings)

5. Limitations

Access to formal models: The lack of access to some formal organizational ontology models prevented a comprehensive analysis and comparison of all available models.

Comparability: Although the foundational models were considered, these models themselves were designed with different objectives, making direct and uniform comparison difficult and influencing the analysis of results.

Model scope: This study focused on foundational models, while purpose-specific models were excluded; therefore, some models were not considered in the analysis.

6. Conclusion

The enterprise ontology offered a comprehensive and systematic framework that enhances the understanding of organizations for both managers and stakeholders. By addressing ambiguities and contradictions, this framework empowered informed decision-making. The clarity it provided is invaluable not only for human users but also for machine processing. This structured understanding is applicable across various domains, including efficient modeling of enterprise architecture, business processes, and context-aware systems.

In this study, we examined and compared the conceptual models of enterprise ontology, highlighting their strengths and weaknesses. Future research should focus on developing a reference enterprise ontology model that encompasses all dimensions of an organization, ensuring it covers every component while identifying key elements and relationships. Moreover, future studies are encouraged to investigate the application of enterprise ontology within specific domains or organizational contexts, in order to evaluate its practical effectiveness and adaptability to particular use cases. Additionally, there is a critical need to establish reliable methods for adapting the reference enterprise ontology model to create tailored models for specific domains, organizations, or applications.

Conflicts of Interests

The authors did not receive support from any organization for the submitted work.

Author Contributions Statement

Seyed Mohsen Rahnamafard conceptualized the research idea, laying the groundwork for the study. Zeinab Rajabi further developed this concept, refining the initial ideas and enhancing the framework for analysis. Both authors, collaborated closely in conducting the analysis and comparisons presented in this study, ensuring a comprehensive evaluation of the existing literature on enterprise ontology.

References

AbdEllatif, M., Farhan, M. S., & Shehata, N. S. (2018). Overcoming business process reengineering obstacles using ontology-based knowledge map methodology. *Future Computing and Informatics Journal, 3*(1), 7-28.

Abramowicz, W., Filipowska, A., Kaczmarek, M., Pedrinaci, C., Starzecka, M., & Walczak, A. (2008). *Organization structure description for the needs of semantic business process management.* Paper presented at the 3rd international Workshop on Semantic Business Process Management colocated with 5th European Semantic Web Conference.

Aguilar, J., Jerez, M., & Rodríguez, T. (2018). CAMeOnto: Context awareness meta ontology modeling. *Applied computing and informatics, 14*(2), 202-213.

Almeida, J. P. A., & Guizzardi, G. (2008). *A semantic foundation for role-related concepts in enterprise modelling.* Paper presented at the Enterprise Distributed Object Computing Conference, 2008. EDOC'08. 12th International IEEE.

Antunes, G., Bakhshandeh, M., Mayer, R., Borbinha, J. L., & Caetano, A. (2014). Using ontologies for enterprise architecture integration and analysis. *CSIMQ, 1*, 1-23.

Aue, A., & Gamon, M. (2005). *Customizing sentiment classifiers to new domains: A case study.* Paper presented at the Proceedings of recent advances in natural language processing (RANLP).

Awadallah, R. (2013). Methods for constructing an opinion network for politically controversial topics.

Brank, J., Grobelnik, M., & Mladenić, D. (2005). A survey of ontology evaluation techniques. In *Proceedings of the conference on data mining and data warehouses (SiKDD 2005)* (pp. 166-170).

Burke, W. W., & Litwin, G. H. (1992). A causal model of organizational performance and change. *Journal of management, 18*(3), 523-545.

Caetano, A., Pereira, C., & Sousa, P. (2012). Generation of business process model views. *Procedia Technology, 5*, 378-387.

Carvalho, V. A., & Almeida, J. P. A. (2015). *A semantic foundation for organizational structures: a multi-level approach.* Paper presented at the 2015 IEEE 19th International Enterprise Distributed Object Computing Conference.

Chen, D., Doumeingts, G., & Vernadat, F. (2008). Architectures for enterprise integration and interoperability: Past, present and future. *Computers in industry, 59*(7), 647-659.

Ciocoiu, M., Nau, D. S., & Gruninger, M. (2001). Ontologies for integrating engineering applications. *Journal of Computing and Information Science in Engineering, 1*(1), 12-22.

Daft, R. L. (2020). *Organization theory & design*: Cengage learning.

Dietz, J. L. (2006). *What is Enterprise Ontology?* : Springer.

DoDAF PLUGIN user guide. (2010). *Version 16.9*.

Fadel, F. G., Fox, M. S., & Gruninger, M. (1994). *A generic enterprise resource ontology.* Paper presented at the Enabling Technologies: Infrastructure for Collaborative Enterprises, 1994. Proceedings., Third Workshop on.

Feilmayr, C., & Wöß, W. (2016). An analysis of ontologies and their success factors for application to business. *Data & Knowledge Engineering, 101*, 1-23.

Fernández-López, M., Gómez-Pérez, A., & Juristo, N. (1997). Methontology: from ontological art towards ontological engineering.

Fox, M. S., Barbuceanu, M., & Gruninger, M. (1995). *An organisation ontology for enterprise modelling: preliminary concepts for linking structure and behaviour.* Paper presented at the Enabling Technologies: Infrastructure for Collaborative Enterprises, 1995., Proceedings of the Fourth Workshop on.

Fox, M. S., & Gruninger, M. (1998). Enterprise modeling. *AI magazine, 19*(3), 109.

Gassen, J. B., Mendling, J., Bouzeghoub, A., Thom, L. H., & de Oliveira, J. P. M. (2017). An experiment on an ontology-based support approach for process modeling. *Information and Software Technology, 83*, 94-115.

Gómez-Pérez, A. (1996). Towards a framework to verify knowledge sharing technology. *Expert Systems with applications, 11*(4), 519-529.

Gordijn, J., Osterwalder, A., & Pigneur, Y. (2005). Comparing two business model ontologies for designing e-business models and value constellations. *BLED 2005 Proceedings*, 15.

Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing? *International journal of human-computer studies, 43*(5-6), 907-928.

Gruninger, M., & Fox, M. S. (1994). An activity ontology for enterprise modelling. *Submitted to AAAI-94, Dept. of Industrial Engineering, University of Toronto, 321*.

Hinkelmann, K., Gerber, A., Karagiannis, D., Thoenssen, B., Van der Merwe, A., & Woitsch, R. (2016). A new paradigm for the continuous alignment of business and IT: Combining enterprise architecture modelling and enterprise ontology. *Computers in Industry, 79*, 77-86.

Hlomani, H., & Stacey, D. (2014). Approaches, methods, metrics, measures, and subjectivity in ontology evaluation: A survey. *Semantic Web Journal, 1*(5), 1-11.

Janićijević, N. (2010). Business processes in organizational diagnosis. *Management: journal of contemporary management issues, 15*(2), 85-106.

Jasper, R., & Uschold, M. (1999). *A framework for understanding and classifying ontology applications.* Paper presented at the Proceedings 12th Int. Workshop on Knowledge Acquisition, Modelling, and Management KAW.

Jung, J. J. (2009). Semantic business process integration based on ontology alignment. *Expert Systems with Applications, 36*(8), 11013-11020.

Kang, D., Lee, J., Choi, S., & Kim, K. (2010). An ontology-based enterprise architecture. *Expert Systems with Applications, 37*(2), 1456-1464.

Kindrick, J. (2009). Enterprise architecture based on design primitives and patterns business transformation agency. *Availaible on: http://www. bta. mil*.

Leppänen, M. (2007). *A context-based enterprise ontology.* Paper presented at the International Conference on Business Information Systems.

Levitt, H. J. (1965). Applied organizational change in industry: Structural, technological, and humanistic approach. *Handbook of Organizations, 1144*, 1170.

McDaniel, M., Storey, V. C., & Sugumaran, V. (2018). Assessing the quality of domain ontologies:

Metrics and an automated ranking system. *Data & Knowledge Engineering, 115*, 32-47.
Mingers, J. (2006). *Realising systems thinking: knowledge and action in management science*: Springer Science & Business Media.
Model, M. The DoDAF Architecture Framework Version 2.02.
Nadler, D. A., & Tushman, M. L. (1995). Types of organizational change: From incremental improvement to discontinuous transformation. *Discontinuous change: Leading organizational transformation*, 15-34.
Officer, C. I. (2009). DoDAF Architecture Framework Version 2.02 - DoD CIO.
O'Leary, D. E. (2010). Enterprise ontologies: Review and an activity theory approach. *International Journal of Accounting Information Systems, 11*(4), 336-352.
OMG. (2008). *Date-Time Foundation Vocabulary Request For Proposal*. Retrieved from???
OMG. (2009). *Organization Structure Metamodel(OSM)*. Retrieved from ?????
OMG. (2015). *Business Motivation Model (BMM)*. Retrieved from http://www.omg.org/spec/BMM/1.3//.
OMG. (2017a). *Unified Architecture Framework (UAF) The Domain Metamodel Version 1.0 - Appendix A*. Retrieved from http://www.omg.org/spec/UAF/1.0.
OMG. (2017b). *Unified Architecture Framework Profile (UAFP)*. Retrieved from http://www.omg.org/spec/UAF/1.0/.
Pateli, A. (2003). A framework for understanding and analysing ebusiness models. *BLED 2003 Proceedings*, 4.
Pereira, D. C. (2015). *Representing organizational structures in enterprise architecture: An ontology-based approach.* Universidade Federal do Espírito Santo.
Pereira, D. C., & Almeida, J. P. A. (2014). *Representing organizational structures in an enterprise architecture language.* Paper presented at the FOMI@ FOIS.
Poels, G., Garcia, F., Ruiz, F., & Piattini, M. (2018). Conceptualizing business process maps. *arXiv e-prints*. Retrieved from https://ui.adsabs.harvard.edu/abs/2018arXiv181205395P.
Porras, J. I., & Robertson, P. J. (1986). *Organization development theory: A typology and evaluation*: Graduate School of Business, Stanford University.
Rajabi, Z., & Abade, M. N. (2012). Data-centric enterprise architecture. *International Journal of Information Engineering and Electronic Business, 4*(4), 53.
Rajabi, Z., Minaei, B., & Seyyedi, M. A. (2013). Enterprise architecture development based on enterprise ontology. *Journal of theoretical and applied electronic commerce research, 8*(2), 85-95.
Rao, L., Mansingh, G., & Osei-Bryson, K.-M. (2012). Building ontology based knowledge maps to assist business process re-engineering. *Decision Support Systems, 52*(3), 577-589.
Robbins, S. P. (1990). *Organization Theory: Structures, Designs, And Applications, 3/e*: Pearson Education India.
Rosing, M. v. (2015). *Overview of the Business Ontology Research & Analysis*.
Saleem, A., & Ghani, U. (2013). Application of Weisbord's organizational diagnosis model: a case of Pakistan banking industry. *Business & Economic Review*, *5*(1), 1-11.
Santos Jr, P. S., Almeida, J. P. A., & Guizzardi, G. (2010). *An ontology-based semantic foundation for ARIS EPCs.* Paper presented at the Proceedings of the 2010 ACM Symposium on Applied Computing.
Santos Jr, P. S., Almeida, J. P. A., & Guizzardi, G. (2013). An ontology-based analysis and semantics for organizational structure modeling in the ARIS method. *Information Systems, 38*(5), 690-708.

Senge, P. M. (1990). The art and practice of the learning organization: New York: Doubleday.

Silva, C., & Belo, O. (2018). *A core ontology for Brazilian higher education institutions.* Paper presented at the CSEDU (2).

Sowa, J. F., & Zachman, J. A. (1992). Extending and formalizing the framework for information systems architecture. *IBM systems journal, 31*(3), 590-616.

Stacey, R. D., Griffin, D., & Shaw, P. (2000). *Complexity and management: Fad or radical challenge to systems thinking?* : Psychology Press.

Thakor, P., & Sasi, S. (2015). Ontology-based sentiment analysis process for social media content. *Procedia Computer Science, 53*, 199-207.

Tham, K. D., Fox, M. S., & Gruninger, M. (1994). *A cost ontology for enterprise modelling.* Paper presented at the Proceedings of 3rd IEEE Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises.

Uschold, M., King, M., Moralee, S., & Zorgios, Y. (1998). The enterprise ontology. *The knowledge engineering review, 13*(1), 31-89.

Villela, K., Santos, G., Schnaider, L., Rocha, A. R., & Travassos, G. H. (2005). The use of an enterprise ontology to support knowledge management in software development environments. *Journal of the Brazilian Computer Society, 11*(2), 45-59.

Waterman Jr, R. H., Peters, T. J., & Phillips, J. R. (1980). Structure is not organization. *Business horizons, 23*(3), 14-26.

Weisbord, M. R. (1976). Organizational diagnosis: Six places to look for trouble with or without a theory. *Group & Organization Studies, 1*(4), 430-447.

Weisman, R. (2011). *An Overview of TOGAF®Version 9.1*. Retrieved from????????

Wierda, G. (2017). *Mastering ArchiMate Edition III: A serious introduction to the ArchiMate enterprise architecture modeling language*: R&A.

Zachman, J. A. (1987). A framework for information systems architecture. *IBM systems journal, 26*(3), 276-292.

Zur Muehlen, M. (2009). Enterprise architecture based on design primitives and patterns-guide for the design and development of event-trace descriptions (DoDAF OV-6c) using BPMN. *Business Transformation Agency of United States Department of Defense, Version, 1*.

Journal of
Knowledge Economy Studies

https://kes.hmu.ac.ir

Hazrat-e Masoumeh University

Readiness Assessment for Big Data Analytics in Citizen Relationship Management: a Case Study of Tehran Governorate

Article Type:
Research Article

Roohallah Noori[1*]
Human Resource Management, Faculty of Management. Kharazmi University, Tehran, Iran
Corresponding Author
E-mail: rnoori@khu.ac.ir

Mojtaba Farrokh[2]
Information Technology and Operation, Faculty of Management. Kharazmi University, Tehran, Iran
E-mail: farrokh@khu.ac.ir

Vahid Heydari[3]
Faculty of Management, Kharazmi University, Tehran, Iran
E-mail: vahid.h1996@gmail.com

Received: 22 December 2025
Revised: 09 February 2026
Accepted: 17 March 2026
Available Online: 18 March 2026

Abstract

This research examined how external organizational factors influence the acceptance and readiness assessment for integrating big data analytics into citizen relationship management (CRM) at Tehran Governorate. The study employed Davis's Technology Acceptance Model (TAM) in an applied, field-based design. Data were collected using a standardized questionnaire based on TAM, with a sample of 105 managers and experts from Tehran Governorate, and analyzed using Structural Equation Modeling (SEM) in SPSS and SmartPLS software. The questionnaire comprised five main dimensions, with validated reliability and validity. Results indicated that external factors (scalability, data storage and processing, data analysis capabilities, flexibility, and reliability), perceived ease of use, and perceived usefulness significantly impact the acceptance of big data analytics. Furthermore, organizational external components such as data storage and processing, flexibility, and reliability lead to satisfaction and intention to utilize big data analytics in managing citizen relations by creating perceived usefulness and confirming user expectations. These findings corroborated the previous research and demonstrated that strategic attention to training, expert recruitment, hardware development, infrastructure enhancement, and information security can facilitate effective adoption of big data analytics, thereby creating opportunities for research development in legal, economic, and other fields.

Keywords

Tehran governorate, Big data, Technology acceptance, Citizen, Citizen relationship management (CRM).

Cite this article: Noori, R., Farrokh, M, & Heydari, V., (2026). **Readiness** assessment for big data analytics in citizen relationship management: a case study of tehran governorate. *Journal of Knowledge Economy Studies (JKES)*, 3(1), 95-120. DOI: http//doi.org/10.22034/kes.2026.2081637.1095

Online ISSN: 3060-7329

Introduction

Countries establish conditions for determining citizenship and rules for its revocation. Citizens, once formally recognized by the state, gain civil, political, and social rights that are less salient for non-citizens. Basic rights include the right to possess a passport, exit and return, and live and work in the country. Some countries permit multiple nationalities, while others emphasize single citizenship (Allen et al., 2020). Corresponding to these rights and responsibilities, urban authorities have reciprocal obligations to citizens, particularly in receiving and addressing public complaints and citizen needs (Ju et al., 2018).

Governments worldwide have targeted improvement of communications with citizens and responsiveness to their requests and complaints. Public sector organizations, by involving citizens in governmental processes, pursue greater transparency. These efforts reflect the importance of effective interaction between government and citizens, with governments leveraging digital transformation to promote citizen participation, social engagement, shared resources, inclusive community consideration, protection of vulnerable groups, and enhanced safety (Chamoso et al., 2020). Such initiatives require societal support and effective communication among all involved groups. Facilitating the citizens' contact with organizations should be prioritized, employing modern tools such as big data analytics, artificial intelligence (AI), and communication programs (Chamoso et al., 2020; Lampropoulos et al., 2022). These technological tools offer significant potential for enhancing service quality and elevating levels of citizen satisfaction.

Citizens view government organizations as experts in various affairs. Whether an organization provides emergency information, facilitates understanding of newly approved laws or policies, or acts as a reliable research resource influences citizen participation and communication patterns (Allen et al., 2020). Increasing citizen knowledge constitutes the foundation for designing digital engagement strategies focused on citizen services. This communications strategy aligns with organizational goals and identifies audiences through digital channels and data analytics. Key avenues to improve citizen engagement through digital communications include identifying audience, branding, sending useful messages, increasing interaction, establishing close outreach, taking person-centered approaches, and conducting precise data analysis (Lampropoulos et al., 2022).

Given these considerations, organizations should effectively apply citizen relationship management (CRM) principles, because citizens within organizations such as provincial governorates are effectively the customers of these organizations. CRM refers to the concepts, tools, and strategies that enhance an organization's ability to understand current and potential customers (Anshari et al., 2018). These systems have been studied across many fields including business, healthcare, science, and other service industries. The widespread adoption of big data has prompted a new perspective in CRM (Anshari et al., 2019; Li et al., 2022).

It is important to distinguish CRM in governmental contexts from traditional business/commercial CRM. While both share foundational principles of managing relationships and leveraging data analytics, governmental CRM fundamentally differs in several key aspects. First, the objective of governmental CRM is public service delivery and democratic accountability rather than profit maximization. Second, citizens are not traditional "customers" who can choose alternatives freely; they are stakeholders with rights and civic responsibilities. Third, governmental CRM must prioritize transparency, data privacy, and equitable service access in ways that exceed commercial requirements. Fourth, success metrics differ. Governmental CRM measures citizen satisfaction, service efficiency, and democratic engagement rather than sales conversion or customer lifetime values. These distinctions necessitate adaptation of CRM principles when applied to public sector organizations such as provincial governorates.

Large volumes of data are exchanged through devices and various systems, collected by search engines and made usable, whether structured or unstructured (Xu et al., 2020). Data from multiple channels are analyzed by organizations to understand patterns and customer behavior. For managing large data volumes in large organizations, big data analytics is essential. It is projected that by 2025 the world will generate 175 zettabytes of data annually, making the traditional processing methods inadequate and necessitating big data analytics (Al-Ateeq et al., 2022). Through using big data, citizens' behaviors, preferences, and needs can be analyzed to improve their experience and increase their satisfaction.

CRM is essentially a process by which an organization manages its interactions with the public and uses data analysis to understand them (Kumar & Reinartz, 2018). Modern CRM systems collect data from diverse communication channels, including corporate websites, telephone, email, live chat, marketing materials, and social media (Naeem et al., 2022). Both public and private sector organizations see big data and its extraction as a major opportunity, with many making substantial investments to collect, integrate, and analyze data (Fallahi Modaresi & Zarei, 2022).

As communications become easier, the data volume related to citizens increases, making analysis more challenging. Here, big data analytics can be effective. With the identified applications of big data analytics, the question arises: do large government agencies—such as ministries and affiliated organizations—possess the capacity and conditions to deploy big data analytics in service of CRM, where the customer is the citizen?

Literature Review

Recent Developments in Adopting Big Data Analytics in Public Sector

Emergent technologies and big data analytics are proving transformative in their impact on organizational performance and public service delivery. Recent empirical evidence demonstrates accelerating adoption of big data analytics in governmental contexts worldwide. The United Kingdom National Health Service achieved more than 90% reduction in patients waiting over 18 months for treatment between September 2021 and May 2023 through the strategic use of big data and analytics (Shukla, 2024). Government departments globally increasingly recognize that becoming data-driven represents primarily a cultural transformation rather than merely a technical upgrade, requiring comprehensive attention to people, processes, and organizational readiness alongside technological infrastructure (Qodea, 2024).

However, government agencies frequently struggle with understaffed IT departments operating outdated legacy systems, making it difficult to adopt big data analytics while ensuring data quality and integrity across thousands of applications (Infosys Public Services, 2024). Political opposition and organizational cultures resistant to change often represent the biggest hindrances to legacy modernization and advanced analytics adoption (Infosys Public Services, 2024). Nevertheless, modern data platforms combined with advancements in AI and generative AI can enhance decision-making processes and improve public services when properly implemented (Shukla, 2024).

Research examining big data analytics in UAE government organizations found that big data analytics capability significantly impacts decision-making capability, with organizational culture and cognitive style of decision makers serving as important mediating factors (Faridoon et al., 2024). This underscores the multifaceted nature of big data adoption, where technical infrastructure alone proves insufficient without appropriate organizational culture and human factors.

Despite concrete evidence of meaningful outcomes, publicly accessible research into big data theory and applications in government sectors remains limited (Hossin et al., 2023). While data is gathered at unprecedented rates, public policy adoption of big data analytics lags due to lack of acceptance and several challenges that limit its utility. The capacity to analyze data has reached unprecedented heights due to IT advancements in both hardware and software, making previously impossible information access now feasible through digitization, AI, computational thinking, and automation (Hossin et al., 2023).

A Model of Technology Acceptance in Contemporary Big Data and AI Contexts

The Technology Acceptance Model (TAM) continues to serve as a robust theoretical framework for understanding technology adoption in the era of big data and AI. A comprehensive 2024 validation study confirmed the TAM's applicability to AI contexts, demonstrating that perceived usefulness remains the most significant predictor of attitude toward use, while perceived ease of use emerged as a significant predictor of both attitude and perceived usefulness (Ibrahim et al., 2024). These findings align with earlier TAM research while extending applicability to advanced analytics technologies.

Recent validation of an extended TAM version in AI contexts, integrating Big Five personality traits and AI mindset factors, confirmed that perceived usefulness is the strongest predictor of attitudes toward AI usage (Ibrahim et al., 2024). Notably, perceived ease of use exerted substantial influence on perceived usefulness, suggesting that easily usable technologies are more likely to be perceived as valuable. Recent applications of TAM have expanded to examine generative AI adoption, finding that simplifying technical complexity enhances the users' perceptions of utility and ease, thereby boosting adoption intentions (Singh, 2024). Key facilitating factors include effective organizational training, intuitive design, and strategic partnerships.

An extended TAM study on AI tool adoption showed that perceived usefulness positively affects attitudes toward using academic AI tools, and these attitudes subsequently predict behavioral intention, which ultimately determines the actual usage behavior (Oubdi & El-Mekkaoui, 2026). Similarly, research on AI tools in recruitment processes indicated that perceived usefulness significantly influences attitude toward using AI, while the relationship between perceived ease of use and behavioral intention was mediated through attitude (Costa et al., 2025). Research on medical students' acceptance of AI technology demonstrated that perceived usefulness, perceived ease of use, and attitude together explained 78% of variance in actual AI use (Asadpoor et al., 2024).

The application of TAM to big data adoption has revealed several critical determinants beyond core constructs. Rahman (2020) identified seven technology-related criteria through conducting a comprehensive analysis. Scalability, data storage and processing capability, performance expectancy, reliability, data analytics capability, flexibility, and output quality functioned as external variables influencing perceived usefulness and ease of use. Research has demonstrated that technological factors, combined with external variables and individual personality traits, positively influence perceived usefulness and perceived ease of use of AI-based technology, while environmental factors such as suggestions from others appeared disruptive to technology acceptance (Na et al., 2022).

Organizational Readiness and Data Governance

Organizational readiness has emerged as a critical factor of success for big data adoption, with data governance serving as a foundational element. Data governance adoption has risen dramatically, with 71% of organizations reporting formal data governance programs in 2024 compared to 60% in 2023 (Galvez, 2024). This surge is driven primarily by the imperative to support AI initiatives, as 62% of organizations identify data governance as the primary challenge inhibiting progress toward AI initiatives (Galvez, 2024).

Organizations with established data governance programs reported substantial benefits including improved data quality (58%), enhanced quality of data analytics and insights (58%), increased collaboration (57%), improved regulatory compliance (50%), and faster access to relevant data (36%) (Pangarkar, 2024). In 2024, more than 65% of data leaders declared data governance as their priority, surpassing AI (44%) and data quality (47%) (Pangarkar, 2024), reflecting strategic recognition that trusted, secure, and compliant data infrastructure is prerequisite to successful AI implementation.

Data governance has shifted from a "nice to have" function to a mainstream necessity, with enterprises increasingly adopting federated governance approaches where central leadership provides frameworks while business units maintain autonomy in implementation (Galvez, 2024). This decentralized-yet-coordinated model addresses the complexity of modern data ecosystems while maintaining organizational agility. Data mesh and data fabric architectures increased from 13% adoption in 2023 to 18% in 2024, further reinforcing the need for robust governance frameworks that support democratized, self-service data access (Pangarkar, 2024).

The year 2024 saw a surge in adoption of AI and machine learning for automating data governance tasks (Smith, 2024). AI's capabilities for pattern recognition and predictive analytics allowed businesses to improve data quality and associated governance processes. The demand for data governance is also driven by data privacy and security concerns, which ranked among the top three priorities for improving data integrity in 2024 (45%) (Galvez, 2024). Regulatory compliance emerged as a goal for 45% of organizations' governance programs, reflecting increased enforcement and severe consequences of non-compliance (Galvez, 2024).

CRM and Smart Cities

The application of big data analytics to CRM represents a convergence of technological capability and democratic governance imperatives. Recent research on digital dialogue in smart cities demonstrated that government response rate, timeliness, and quality significantly impact citizen satisfaction, with communication satisfaction, timeliness satisfaction, and resolution satisfaction all influenced by government responsiveness (Cao & Kang, 2025). This empirical finding underscores the importance of not merely collecting citizen data but responding effectively to citizen concerns.

CRM systems have been proven to be effective in initiating smart city services, enhancing public policy implementation, and improving government representativeness across multiple international contexts, including the 311 system in the United States, China's 12,345 hotline, and Taiwan's 1999 system (Cao & Kang, 2025). These systems leverage big data analytics to transform how governments interact with citizens and respond to public needs. However, implementation of CRM systems for public services can introduce complexity and political dimensions, as these initiatives are frequently used in election rhetoric or become topics of political debate, potentially limiting development and local-level applications (Cao & Kang, 2025).

Modern CRM solutions facilitate two-way communications through notifications, automated responses, and real-time updates, while serving as data management hubs that provide valuable insights into citizen engagement patterns and feedback trends (Rattletech 2021; Civita, 2025). This data-driven approach enables city officials to make informed decisions and adjust strategies based on evidence rather than intuition. Recent smart city governance research emphasized that active citizen engagement remains a key focus, with collaborative environments driven by, for, and with citizens seeking to build smart cities embedded in local realities (Baycan & Yigitcanlar, 2024).

Traditional barriers to citizen engagement persist, with town hall meetings and conventional surveys leaving citizen participation rates low and a majority of citizens uninformed about municipal activities (Beesmart City, 2025). Digital platforms that function 24/7, provide multiple access channels, and ensure data security are increasingly essential for meeting citizen expectations and fostering trust (Beesmart City, 2024). Transparency in data usage and demonstrable service improvements based on citizen input was proved to be critical for building and maintaining public confidence in government digital initiatives (Civita App, 2025).

Infrastructure and Technological Determinants

Prior studies underscored the critical importance of technological infrastructure in big data adoption. Fallahi Modaresi and Zarei (2022) demonstrated that understanding relative advantage and addressing technical readiness are key factors in adopting new technologies in tourism industry, while compatibility may have less significant impact than traditionally assumed. Noori et al. (2017) found that job position and organizational support structures significantly influence information technology adoption in human resource management contexts, suggesting that hierarchical and cultural factors mediate technology acceptance.

Alyusuf and Al-Rahmi (2022) applied TAM to explore the adoption of big data analytics in higher education, revealing that facilitating conditions and perceived risk are important determinants of attitudes toward use. Their findings emphasized that organizational infrastructure—including technical support, training programs, and risk mitigation strategies—directly shapes user perceptions and adoption intentions. Ghali et al. (2021) assessed big data adoption readiness using a Technology-Organization-Environment (TOE) framework, identifying the interdependence of technological capabilities, organizational culture, and environmental pressures in determining the success of adoption.

More recent research by Al-Dossari et al. (2023) examined factors influencing the adoption of big data analytics through systematic literature review, identifying executive support, training, and robust governance frameworks as key enabling factors. Their research emphasized that successful adoption requires coordinated attention to multiple organizational dimensions rather than focusing solely on technical infrastructure. Research has also investigated factors influencing smart city development using an integrated approach of big data technologies, Internet of Things (IoT), and cloud computing, demonstrating that technological integration must align with organizational readiness and strategic vision (Fahm Fāmm & Hamidi, 2018).

Research Gap

While the existing research has examined technology adoption across various contexts and explored the applications of big data in both private and public sectors, a significant gap remains in understanding how large governmental institutions—specifically provincial governorates—can assess and enhance their readiness to implement big data analytics for CRM. Previous studies have primarily focused either on technical aspects of adopting big data (e.g., Al-Dossari et al., 2023; Ghali et al., 2021) or on general technology acceptance in service organizations (e.g., Noori & Emamvirdi, 2015), but few have integrated these perspectives within the specific operational and political context of citizen-facing governmental agencies.

Moreover, while recent research has validated TAM in AI and advanced analytics contexts (e.g., Costa et al., 2025; Ibrahim et al., 2024), these studies have predominantly examined commercial organizations or educational institutions. The unique challenges facing large public sector organizations—including bureaucratic structures, political considerations, budget constraints, and complex stakeholder environments—require dedicated investigations. The integration of big data analytics into governmental CRM introduces additional complexities related to data privacy, transparency requirements, and democratic accountability that are less prominent in private sector contexts.

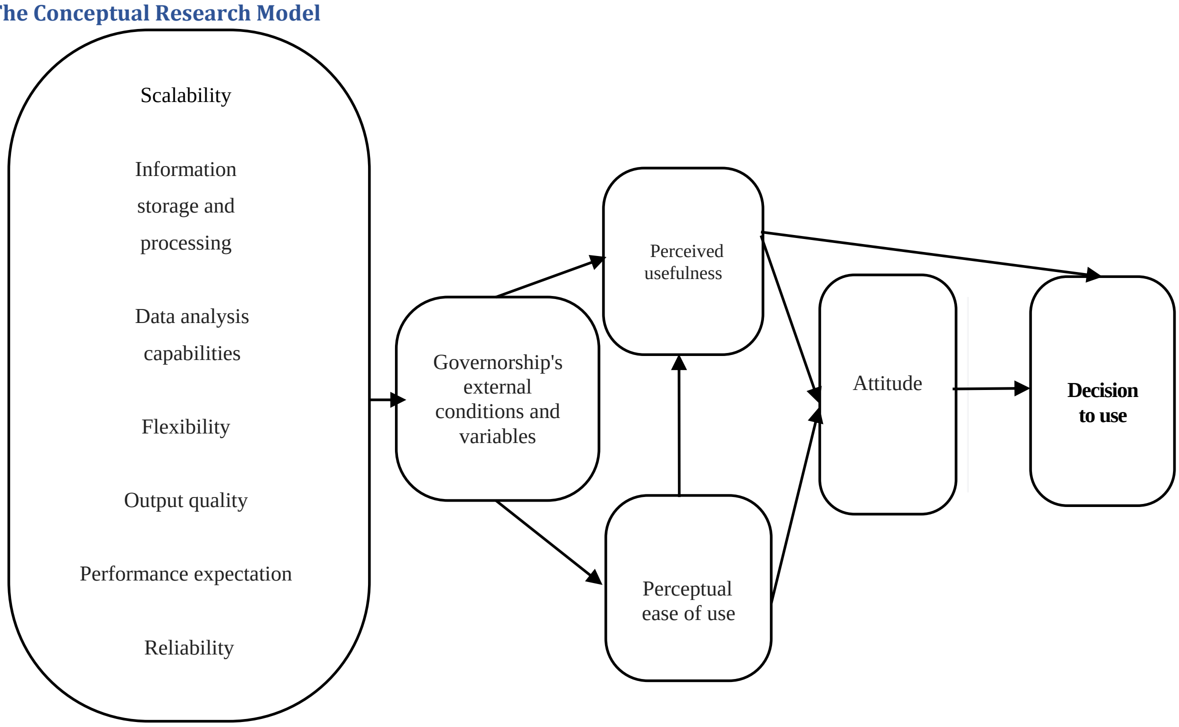
This study addressed these gaps by applying Davis's TAM to assess Tehran Province Governorate's readiness to adopt big data analytics for CRM. The research examined the interplay between external organizational factors (including scalability, data processing capabilities, flexibility, and reliability) and psychological acceptance factors (perceived usefulness, perceived ease of use, attitude, and behavioral intention). The study contributed to empirical evidence on how infrastructure capabilities and user perceptions interact in determining big data adoption readiness in a large-scale public sector context, providing insights relevant to similar governmental organizations facing digital transformation challenges worldwide.

Conceptual Model and Hypotheses

The Conceptual Framework

Based on Davis's TAM and the reviewed literature, this study proposed a conceptual model (Fig.1.) that integrates external organizational factors with core TAM constructs. The model regards assessing environmental conditions and initial needs as essential in the first stage of technology acceptance. Researchers have emphasized that for introducing new tools like big data analytics, it is necessary to examine infrastructure, drivers of use, and costs (Aldholay et al., 2022).

Figure 1.
The Conceptual Research Model



(Source: The Researcher's Findings)

The proposed model incorporates seven technology-related criteria identified by Rahman (2020):

- Scalability: The ability of software and hardware to handle increasing workloads and data;
- Data storage and processing capability: The ability to store and process hundreds of terabytes of data with modern models;
- Performance expectancy: The effectiveness of the technology in use and its importance for technology adoption;

– Reliability: Providing greater reliability while preserving data integrity across multiple nodes;
– Data analytics capability: The ability to uncover patterns and utilize data mining and machine learning;
– Flexibility: The possibility of extracting and processing data from diverse sources, both structured and unstructured;
– Output quality: The importance of preserving data quality and meaningful relationships for accurate decision-making.

Each of these external organizational factors represents a critical dimension of technological readiness and is theoretically linked to TAM's core constructs as follows: Scalability refers to the system's capacity to handle growth in data volume, user numbers, and analytical complexity without performance degradation. This factor directly influences the perceived ease of use. When systems scale seamlessly, users experience consistent performance regardless of workload, reducing friction and enhancing usability perceptions (Rahman, 2020). Data storage and processing capability encompasses the infrastructure's ability to store and process hundreds of terabytes of structured and unstructured data using modern distributed computing architectures. This capability is foundational to perceived usefulness because without adequate storage and processing power, big data analytics cannot deliver meaningful insights (Mikalef et al., 2020). Data analytics capability represents the system's ability to apply advanced techniques including data mining, machine learning, pattern recognition, and predictive modeling to extract actionable insights. This factor most directly impacts perceived usefulness (Maroufkhani et al., 2020). Flexibility indicates the system's capacity to integrate and process data from diverse sources in various formats. Flexible systems enhance both perceived usefulness and perceived ease of use (Rahman, 2020). Reliability ensures data integrity, system availability, and consistent performance across distributed nodes. High reliability builds user trust and confidence, directly influencing perceived usefulness (Ghali et al., 2021). These external factors collectively shape the technological environment within which users form perceptions about adopting big data analytics.

Big data analytics, together with IoT and cloud infrastructure, play a significant role in CRM, especially in smart cities. Key applications include smart cities and contactless technologies, various contactless programs leveraging big data analytics in public services, transportation, healthcare, and emergency management, using IoT sensors and environmental monitoring to collect and analyze data in real time, and assisting in effective management of emergency situations (Al Batayneh et al., 2021; Ashiku et al., 2021; Cabrera-Sánchez & Villarejo-Ramos, 2020).

Research Hypotheses

Davis's TAM provides a good fit for assessing the readiness of Tehran Province Governorate to use big data analytics for CRM.

Hypothesis 1: External variables (organizational conditions) at Tehran Province Governorate have a meaningful relationship with perceived usefulness of using big data.

The relationship between perceived usefulness of big data usage and internal organizational conditions is of substantial importance. Big data refers to large collections of data gathered from diverse sources such as sensors, social media, and transactional systems (Maroufkhani et al., 2020). Using big data analytics enables organizations to exploit vast data, extract patterns, relationships, and conceptual information (Mikalef et al., 2020). Organizational factors such as culture, staff capabilities, policies, and leadership significantly influence big data usage.

Hypothesis 2: Organizational conditions have a significant relationship with perceived ease of using big data.

Perceived ease of use indicates the extent to which users can interact with tools with ease (Mikalef et al., 2020). If users find big data tools simple and efficient, they are likely to use them for decision-making and analysis. Key organizational factors that can help improve perceived usefulness include relative advantage perception, organizational readiness, top management support, government regulation, and link to firm performance (Caffaro et al., 2020).

Hypothesis 3: Perceived ease of using big data has a significant relationship with perceived usefulness.

The Technology Acceptance Model suggests that when a technology is user-friendly, users find it more useful and hence use it more (Siagian et al., 2022). In the context of big data, user-friendly tools help people find data useful and make informed decisions. If system use is easy, it is perceived as more useful, especially given the volume, speed, and variety of data in big data contexts.

Hypothesis 4: Perceived usefulness of using big data has a significant relationship with attitude toward using big data analytics.

Attitude is recognized as a key variable in technology acceptance and exerts a strong influence on willingness to use technologic services. A positive attitude of citizens toward big data can lead to improved services (Kumar & Reinartz, 2018). If users regard the technology as useful, they develop a positive attitude that leads to greater acceptance and effective use of big data analytics (Arghashi & Yuksel, 2022).

Hypothesis 5: Perceived usefulness of using big data has a significant relationship with the decision to use big data analytics.

The perceived usefulness of big data plays an important role in decisions to employ data analytics (Kar & Dwivedi, 2020). The decision to use big data analytics is influenced by the perceived benefits that these analyses bring (Sarker, 2021). Training staff by senior management can improve employees' understanding of the technology and facilitate the decision to use it. To encourage use of big data analytics, its usefulness should be demonstrated concretely. For example by presenting successful case studies, training on potential applications, and ensuring user-friendliness of analytic tools (Xu et al., 2020).

Hypothesis 6: Perceived ease of using (PEOU) big data analytics tools is related to attitudes toward using them.

This relationship is essential for user acceptance and successful implementation of big data technologies. The attitude toward using big data analytics includes users' affective responses, encompassing perceived ease of use, perceived usefulness, and understanding of how to use the technology (Islami et al., 2021). When big data analytics tools are designed with the user in mind and assured of simplicity and ease of use, they are more likely to be adopted. To enhance attitude and better understanding of ease of use, employing a range of supports—such as providing help resources, training, providing a test-friendly environment and opportunities for hands-on experience, networking with peer organizations, and fostering a culture of innovation—is essential (Iriani & Andjarwati, 2020).

Hypothesis 7: Attitude toward using big data is significantly related to the intention to use big data.

A positive attitude toward big data is often a strong predictor of the intention to actually use big data analytics in practice (Kasilingam, 2020). This attitude reflects a general evaluation of data analytics, and if positive, it reinforces beliefs about the improved performance, insights, and outcomes. The intention to use big data results from attitude, where a positive attitude can lead to higher intention to use. Habits of using up-to-date tools can also influence attitude and adoption of new technologies (Aworh et al., 2021). To cultivate a positive attitude toward big data, it is crucial to demonstrate benefits, share success stories, ensure ease of use, and provide support and training to potential users.

Method

This study adopted a pragmatic research approach employing survey methodology for primary data collection. Given the nature of the conceptual model, the research approach is descriptive and of a correlational branch. To collect information, the study employed both library (documentary) and field methods. Initially, a literature review encompassed a range of books, articles, and specialized theses related to the research domain.

The study population comprised managers and staff from the Public Communications and Information Technology departments, along with subject matter experts in the research domain. A sample of 105 respondents was drawn based on staff numbers without prior acquaintance. Data were collected using a standardized questionnaire based on Davis's TAM. This questionnaire comprises five main dimensions, and its validity and reliability were confirmed by Randi et al. (2014). Cronbach's alpha for the questionnaire was approximately 0.78, indicating an acceptable level of reliability. Composite reliability and extracted average variance (AVE) were also above 0.75 and 0.5, respectively, demonstrating the reliability of the model.

While Rahman's (2020) original framework includes performance expectancy as a distinct construct, we operationalized this concept through the perceived usefulness dimension of TAM. The performance expectancy overlaps substantially with Davis's perceived usefulness construct. To avoid conceptual redundancy and maintain parsimony, we incorporated performance-related items within the perceived usefulness scale. Future research may benefit from explicitly distinguishing these constructs to examine their unique contributions.

In this study, Cronbach's alpha, composite reliability, and construct validity were calculated and reported. In the section on organizational characteristics, the extracted factors are presented and displayed in the conceptual model. Descriptive statistics include frequency, mean, and standard deviation. Normality of the variables was assessed via Skewness and Kurtosis. Pearson correlations and the validity and reliability of the questionnaire were examined using Confirmatory Factor Analysis. The theoretical model was tested with (SEM). SPSS and SmartPLS software were used for data analysis. The significance level (alpha) for hypothesis testing was set at 0.05.

Findings

Demographic Characteristics

Based on the collected data, the demographic characteristics are presented in Table 1.

Table 1.
Description of Demographic Variables with Frequency Counts and Percentages

Percent	Number	Classification	Variable
68.16	72	male	Gender
31. 4	33	female	
27.6	29	below the age of 30	Age
38.1	40	between 30 and 40 years old	
34.3	36	aged more than 40	
8.6	9	High school diploma	Education level
14.3	15	Associate degree	
49.5	52	Bachelor's degree	
27.6	29	Master's degree or higher	
43.8	46	Humanities	Field of study
21.0	22	Experimental Sciences	
19.0	20	Mathematics and Physics	
16.2	17	Other fields	
22.9	24	Less than 5 years	Activity history
30.5	32	5 to 10 years	
46.7	49	More than 10 years	
11.4	12	Very low	Computer familiarity
27.6	29	Low	
37.1	39	Adequate	
20.0	21	High	
3.9	4	Professional	
3.8	4	Very low	Internet familiarity
8.6	9	Low	
45.7	48	Adequate	
30.5	32	High	
11.4	12	Professional	

(Source: The Researcher's Findings)

The results indicated that 68.6% of respondents were male and 31.4% were female. The largest age group (38.1%) was between 30 and 40 years old. The bachelor's degree holders represented the largest educational category at 49.5%. The majority of respondents (43.8%) reported Humanities as their field of study. The largest share of respondents (46.7%) had more than 10 years of work experience. Regarding the computer familiarity, 37.1% had adequate familiarity, while 45.7% had adequate Internet familiarity.

Descriptive Statistics

Table 2.
Descriptive Statistics of Core Variables

Descriptive Statistics				Variables
Maximum Score	Minimum Score	Standard Deviation	Mean	
4.40	2.00	0.64	3.08	External Factors
5.00	1.83	0.84	3.33	Perceived Usefulness
4.75	1.50	0.72	2.89	Perceived Ease of Use
5.00	2.40	0.74	3.60	Attitude
5.00	2.67	0.65	3.65	Behavioral Intention to Use

(Source: The Researcher's Findings)

Table 2 revealed that the mean score for the "perceived ease of use" was below the theoretical average of 3, whereas all other variables exhibited mean scores above this threshold. The highest mean was observed for the "intention to use" scale, with a value of 3.65. The assumptions underlying SEM—including data distribution and absence of multicollinearity—were examined and confirmed, validating the suitability of the dataset for conducting SEM analysis.

Assessing Validity and Reliability

Table 3.
The Results of Confirmatory Factor Analysis

Cronbach's Alpha	Composite Reliability	AVE	Variables
0.90	0.92	0.66	Perceived Usefulness
0.81	0.87	0.63	Perceived Ease of Use
0.86	0.97	0.64	Attitude
0.61	0.79	0.56	Behavioral Intention to Use
0.75	0.81	0.47	External Factors

(Source: The Researcher's Findings)

The results of validity and reliability tests indicated that all observed variables demonstrated acceptable construct validity. All factor loadings had t-values greater than 1.96, indicating statistical significance at a minimum confidence level of 95% ($p < 0.05$). Two items from the "external factors" construct—items 23 and 24—were excluded from the analysis due to factor loadings below 0.50 and weak validity and reliability.

The reliability analysis showed that composite reliability values ranged from 0.79 for "intention to use" to 0.92 for "perceived usefulness," confirming reliability via the composite reliability method. The Cronbach's alpha values ranged from 0.61 for "intention to use" to 0.90 for "perceived usefulness," indicating acceptable internal consistency. Given the novelty of the questionnaire and the limited number of items (only three) in the "intention to use" scale—and considering that Cronbach's alpha is sensitive to item count—values above 0.60 were deemed acceptable.

The convergent validity was assessed using the Average Variance Extracted (AVE) index. All AVE values were close to or exceeded the recommended threshold of 0.50. The lowest AVE value was 0.47 for the "external factors" scale, while the highest was 0.66 for "perceived usefulness," confirming the convergent validity. Discriminant validity was evaluated using both the Fornell-Larcker criterion and the HTMT ratio. Results from both methods supported the presence of discriminant validity. HTMT values for all constructs were below the 0.85 threshold, confirming the discriminant validity.

Correlation Analysis

Table 4.
The Pearson Correlation Matrix among Core Research Variables**

Behavioral Intention to Use	Attitude	Perceived Ease of Use	Perceived Usefulness	External Factors	Variables
				1	External Factors
			1	**0.41	Perceived Usefulness
		1	**0.49	**0.33	Perceived Ease of Use
	1	0.09	**0.37	**0.41	Attitude
1	0.62	*0.19	**0.39	**0.52	Behavioral Intention to Use

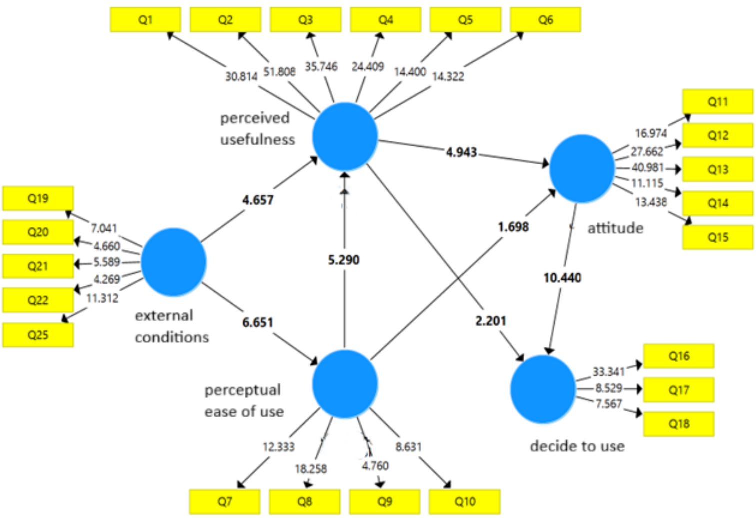
(Source: The Researcher's Findings)

The results indicated statistically significant relationships between the four predictor variables—external factors, perceived usefulness, perceived ease of use, and attitude—and the variable "intention to use" ($p< 0.05$). All four predictors were positively correlated with intention to use. The strongest association with intention to use was observed for attitude ($r= 0.62$), followed by external factors ($r= 0.52$), perceived usefulness ($r = 0.39$), and perceived ease of use ($r= 0.19$).

The Results of SEM

The conceptual model was tested using SEM based on Partial Least Squares (PLS). The analysis was conducted using SmartPLS software. Due to violation of the multivariate normality assumption, the PLS method was selected as an appropriate alternative.

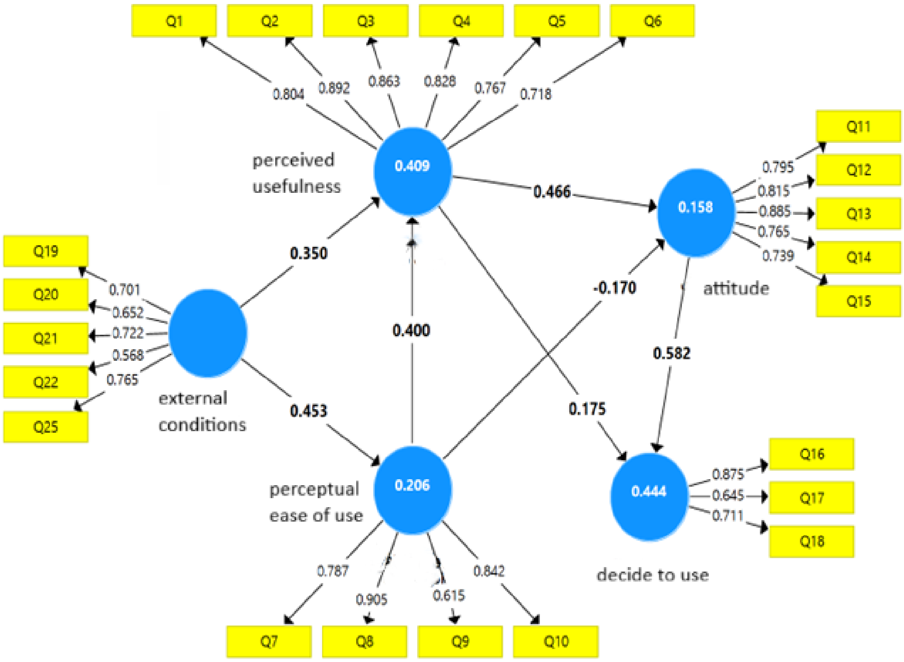
Figure 2.
The Research Model in T-value Mode



(Source: The Researcher's Findings)

Figure 2 presents the empirical model in terms of t-values. Statistically, if the absolute value of a t-statistic exceeds 1.96, the corresponding relationship is considered significant at a confidence level of at least 95% ($p < 0.05$). The findings indicated that out of seven paths or relationships in the model, six were statistically confirmed.

Figure 3.
The Research Model in standard Path Coefficients Mode



(Source: The Researcher's Findings)

Figure 3 presents the model in terms of standardized coefficients. The analysis of t-values indicated that out of seven paths in the model, six were statistically confirmed ($|t| > 1.96$, $p < 0.05$). The analysis of standardized coefficients revealed that the strongest effect within the model was the influence of attitude on intention to use, with a coefficient of 0.58. The second strongest effect was the impact of perceived usefulness on attitude, with a coefficient of 0.46.

The model's coefficient of determination (R^2) was 0.44, indicating that the predictor variables accounted for 44% of the variance in intention to use. This reflects reasonably strong explanatory power of the model.

Table 5.
The Research Model Fit Indices**

Standardized Root Mean Square Residual (SRMR)	Normed Fit Index (NFI)	Predictive Relevance Index (Q^2)	Coefficient of Determination (R^2)	Dependent Variable
0.166	0.32	0.23	0.44	Intention to Use

(Source: The Researcher's Findings)

According to Table 5, the Normed Fit Index (NFI) was calculated at 0.32, which falls below the acceptable threshold of 0.90. The Standardized Root Mean Square Residual (SRMR) was 0.166, exceeding the recommended threshold of 0.10. The predictive relevance (Q^2) for the construct of "intention to use" was 0.23, which lies between moderate and strong levels, suggesting that the model demonstrates acceptable predictive capability. The coefficient of determination (R^2) for "intention to use" was 0.44, exceeding the moderate benchmark of 0.33, reflecting satisfactory explanatory power.

The Explanation of the Model Fit Indices: The model fit indices revealed mixed results that warrant discussion. While the coefficient of determination (R^2 = 0.44) and predictive relevance (Q^2= 0.23) demonstrated acceptable to moderate levels, the Normed Fit Index (NFI= 0.32) and Standardized Root Mean Square Residual (SRMR = 0.166) fell below the conventional thresholds. Several factors may explain these weaker fit indices. First, the complexity of adopting big data in governmental contexts may involve additional unmeasured variables such as political factors, inter-departmental coordination, and bureaucratic processes not captured in the standard TAM framework. Second, the exclusion of two items from the construct of external factors due to low factor loadings may have affected the overall model fit. Third, the relatively small sample size (n=105) and the specific organizational context (a single governorate) may limit the model generalizability. Despite these limitations, the significant path coefficients and moderate explanatory power suggested that the model provides valuable insights into adoption dynamics. Future research should consider expanding the model to include context-specific variables relevant to public sector organizations and employ larger, multi-organizational samples to improve the model fit.

The Results of Hypothesis Testing

Table 6.
The Results of Hypothesis Testing

Status	p-value	t-value	Standard Error	Standardized Coefficient	type of relationship
confirmation	0.001<	4.66	0.07	0.35	← External Factors Perceived Usefulness
confirmation	0.001<	6.65	0.06	0.45	← External Factors Perceived Ease of Use
confirmation	0.001<	5.29	0.07	0.40	← Perceived Ease of Use Perceived Usefulness
confirmation	0.001<	4.94	0.09	0.46	← Perceived Usefulness Attitude
confirmation	0.028	2.20	0.07	0.17	← Perceived Usefulness Behavioral Intention to Use
Not confirmation	0.090	1.70	0.10	0.17-	← Perceived Ease of Use Attitude
confirmation	<0.001	10.44	0.05	0.58	← Attitude Behavioral Intention to Use

Sensitivity Analysis

To assess the robustness of our findings, we conducted several sensitivity analyses examining the model stability and potential biases. We re-estimated the model separately for key demographic subgroups to assess the stability of path coefficients. Results showed consistent patterns across education levels (bachelor's vs. master's and higher), age groups (under 40 vs. 40 and over), and work experience categories (under 10 years vs. 10+ years). The path coefficients varied by less than 0.10 across subgroups, and the pattern of significance remained consistent, indicating the model stability across respondent characteristics. Examination of standardized residuals identified three potential outliers (standardized residuals > ±3.0). Re-estimating the model after excluding these cases resulted in minimal changes to path coefficients (maximum change: 0.04) and no changes in hypothesis testing conclusions, confirming that results were not driven by extreme values. While we used PLS-SEM due to non-normality, we also estimated the model using covariance-based SEM (CB-SEM) with maximum likelihood estimation for comparison. Despite the normality violation, CB-SEM results showed similar path coefficients (correlations > 0.90 between PLS and CB estimates) and identical patterns of significance, providing convergent validity for our findings. We also conducted Harman's single-factor test to assess the potential common method bias. Exploratory Factor Analysis with all items showed that the first unrotated factor explained 32% of variance, well below the 50% threshold that would indicate substantial common method bias. Additionally, we examined the correlation matrix. The highest correlation between constructs was 0.62 (Attitude-Intention), below the 0.90 threshold suggesting method bias concerns. These results indicated that the common method bias is unlikely to substantially affect our findings. Collectively, these sensitivity analyses demonstrated that our findings are robust across analytical choices and not substantially influenced by outliers, demographic heterogeneity, or methodological artifacts.

Discussion

The findings indicated that external factors—such as technological advancements, system integration, and improved analytical output quality—have a significant impact on the perceived usefulness of big data utilization within Tehran Governorate. Advanced technologies enhance data collection and analysis capabilities, leading to more informed and innovative decision-making. These results align with studies examining the influence of technologies and organizational conditions on the adoption of emerging technologies (e.g., Al-Dossari et al., 2023; Ghali et al., 2021). Similarly, research has confirmed the role of infrastructure and big data analytics capabilities in project success (e.g., Faridoon et al., 2024; Shukla, 2024).

The results also showed that external factors—such as scalability, enhanced processing, big data capabilities, and system performance—significantly affect the perceived ease of using big data technologies in organizations. These factors contribute to reduced analysis time and improved user experience. These findings are consistent with studies confirming the influence of external factors on ease of technology use (Noori, A., Hatami, Z., & Ebrahimiān, H., 2017; Hamta, N., Mohammadzadeh, Y., Hemati, M., & Dehghanzadeh, R., 2020).

The findings revealed that perceived ease of use has a significant impact on perceived usefulness. When users can easily access and interact with big data systems, they are more likely to recognize the value and utility of the data. The results demonstrated that perceived usefulness of big data significantly influences users' attitudes toward its adoption. A positive perception of data utility can enhance the user trust and foster a more favorable overall attitude. This finding is consistent with studies confirming the impact of perceived usefulness on user attitudes (Barham, H., & Daim, T., 2020; Fallahi Modaresi, S., & Zarei, A., 2022).

Furthermore, the findings indicated that perceived usefulness significantly affects the users' decisions to adopt big data technologies. A positive perception of data utility increases trust and willingness to integrate big data into decision-making processes. This result is supported by studies confirming the role of perceived usefulness in adoption decisions (Noori, A., Hatami, Z., & Ebrahimiān, H., 2017; Hamta, N., Mohammadzadeh, Y., Hemati, M., & Dehghanzadeh, R., 2020).

The results showed that perceived ease of use does not significantly influence attitudes toward big data analytics. This suggests that ease of system interaction alone is insufficient to foster positive attitudes. Additional factors such as training and organizational culture are also required. A similar conclusion was reached in earlier research (e.g., Noori & Emamvirdi, 2015).

The non-significant relationship between perceived ease of use and attitude toward big data analytics (H6: $\beta = -0.170$, $p = 0.090$) warrants careful examination, given the central role of PEOU in the original TAM. From a theoretical perspective, this finding may reflect the complexity of adopting big data analytics in governmental contexts, where the perceived strategic value outweighs the operational convenience. Big data analytics represents sophisticated technology where users—especially senior managers and experts—may prioritize outcomes over ease of operation. Our findings support an alternative pathway: PEOU significantly influences perceived usefulness (H3: $\beta = 0.400$, $p < 0.001$), which in turn strongly affects attitude (H4: $\beta = 0.466$, $p < 0.001$). This suggests that ease of use indirectly shapes attitudes by first enhancing the users' beliefs about the system's utility. Contextually, the sample comprised primarily managers and technical experts with relatively high digital literacy. For technically sophisticated users, ease of use may be less salient in attitude formation because they already possess competencies to navigate complex systems. The finding aligns with previous research in public sector contexts where perceived usefulness emerged as a stronger predictor of technology acceptance than ease of use (Noori & Emamvirdi, 2015). From a practical standpoint, this suggests that organizations should prioritize demonstrating concrete values and strategic benefits rather than solely focusing on interface simplification. Training programs should emphasize use cases, success stories, and organizational benefits alongside technical skills development.

Additionally, the findings confirmed that user attitudes toward big data significantly influence their decision to adopt these technologies. A positive attitude facilitates faster and more effective acceptance within Tehran Governorate. This result aligns with studies validating the impact of attitude on decision-making.

The successful implementation of big data analytics in citizen relationship management requires appropriate infrastructure, skilled human resources, digital culture development, information security, and digital governance. These elements contribute to user trust and acceptance. Overall, this study demonstrated that big data analytics can enhance the efficiency and effectiveness of public services. By improving key factors, organizations can achieve more responsive and efficient citizen engagement.

Conclusions and Recommendations

This research examined the readiness of Tehran Province Governorate to adopt big data analytics for CRM using Davis's TAM. The findings provided valuable insights specific to this organizational context, with potential relevance to similar governmental organizations when appropriately adapted.

Findings Specific to Tehran Province Governorate: Based on our analysis of 105 managers and experts from Tehran Province Governorate, we found that external organizational factors (scalability, data storage and processing, flexibility, and reliability) significantly impact both perceived usefulness and perceived ease of using big data analytics. These findings underscored the critical importance of technological infrastructure at Tehran Governorate specifically. Current IT systems must be upgraded to support scalable architectures, enhanced storage capacity, flexible data integration, and robust reliability to facilitate the user acceptance. The study validated core TAM relationships in Tehran Governorate's context. Perceived usefulness emerged as a strong predictor of both attitudes and behavioral intentions among Governorate personnel, while perceived ease of use significantly influenced perceived usefulness. Notably, perceived ease of use did not directly impact attitude in this context, which may reflect the technical sophistication of our sample (managers and experts) and the strategic nature of big data analytics deployment in governmental settings. Attitude toward using big data analytics proved to be the strongest predictor of behavioral intention ($\beta = 0.582$) among Tehran Governorate respondents, highlighting the critical importance of fostering positive attitudes through demonstration of concrete benefits, sharing success stories, and creating supportive organizational culture within the Governorate.

Practical Implications for Tehran Province Governorate: Based on our findings, we propose the following recommendations specifically tailored to Tehran Province Governorate's current situation:

Infrastructure Development Priorities:

- Immediate: Assess and upgrade current data storage infrastructure to handle projected citizen data growth. Tehran Governorate currently manages citizen interactions through multiple channels (in-person offices, telephone hotlines, website contact forms, and social media). These channels generate fragmented data that requires unified storage and processing capability.
- Short-term (6-12 months): Implement scalable cloud-based or hybrid infrastructure capable of processing large volumes of citizen complaints, requests, and feedback data. Current systems likely cannot handle advanced analytics at scale.
- Medium-term (1-2 years): Deploy big data analytics platforms with flexible data integration capabilities to consolidate data from all citizen touchpoints, enabling comprehensive relationship management.

Specific Actions of Human Resource Development:

- Conduct needs assessment of current analytical skills among IT and public communication departments at Tehran Governorate.
- Design and implement targeted training programs focusing on big data analytics tools, emphasizing practical applications to CRM (e.g., complaint pattern analysis, service demand forecasting, citizen sentiment analysis).
- Recruit specialized data scientists and analysts with expertise in public sector analytics.
- Establish mentorship programs pairing technically sophisticated staff with that requiring skill development.

Demonstrating Concrete Values to Build Positive Attitudes: Given the strong role of attitude in predicting intention ($\beta = 0.582$), Tehran Governorate should prioritize initiatives that foster positive perceptions:

- Implement pilot projects analyzing citizen complaint patterns to identify service gaps and demonstrate tangible improvements.
- Showcase quick wins such as reduced response times to citizen inquiries through analytics-driven resource allocation.
- Develop dashboards visualizing citizen satisfaction trends and service performance metrics accessible to managers.
- Share success stories internally highlighting how data insights have improved decision-making Organizational Culture and Change Management:
- Secure visible commitment from senior leadership at Governorate level, with explicit endorsement of data-driven CRM.
- Establish cross-departmental working groups to break down data silos between IT, public relations, and service delivery departments.
- Create incentive structures rewarding data-driven decision-making and citizen-centric service innovations.
- Communicate the strategic importance of big data analytics through regular internal communications.

Data Governance and Citizen Trust:

- Establish clear data governance policies addressing citizen data privacy, security, and ethical use—critical for maintaining public trust- Implement robust cyber-security measures protecting sensitive citizen information.
- Develop transparent communications to citizens explaining how their data improves services while protecting privacy.
- Ensure compliance with relevant Iranian data protection regulations.

Phased Implementation Approach for Tehran Governorate:

Phase 1 (Months 1-6): Infrastructure assessment, pilot project selection, training needs analysis.

Phase 2 (Months 7-12): Infrastructure upgrades, initial training programs, first pilot project implementation.

Phase 3 (Year 2): Full-scale deployment, continuous improvement, expansion to additional use cases.

Expected Benefits Specific to Tehran Governorate Context:

- More responsive citizen services through rapid identification of emerging issues and concerns - Evidence-based resource allocation across Governorate districts based on citizen demand patterns.
- Improved transparency and accountability through data-driven performance monitoring - Enhanced ability to anticipate and address citizen needs proactively.
- Strengthened legitimacy and public trust through demonstrably improved services.

Limitations and Generalizability Considerations

This study has several important limitations that affect its generalizability. First, findings are based on a single organization (Tehran Province Governorate) at a single point in time. While insights may be relevant to other Iranian provincial governorates or similar governmental organizations in other countries, direct generalization requires caution. Organizational culture, technological maturity, budget constraints, political contexts, and citizen demographics vary substantially across jurisdictions. Second, our sample of 105 managers and experts, while appropriate for this organization, may not represent perspectives of front-line staff that interact directly with citizens daily. Their attitudes and perceptions may differ significantly. Third, the cross-sectional design captures intentions rather than actual adoption behavior. Longitudinal research tracking Tehran Governorate through implementation would provide more definitive evidence of actual adoption and sustained use. Fourth, the weaker model fit indices suggest that additional factors beyond our theoretical framework may influence adoption in governmental contexts. Political factors, inter-agency coordination requirements, budget approval processes, and public accountability mechanisms likely play important roles not captured by standard TAM.

Recommendations for Other Governmental Organizations

While our findings are most directly applicable to Tehran Province Governorate, similar governmental organizations considering big data analytics for CRM may find value in our results, with appropriate contextual adaptation. Organizations should:

- Conduct organization-specific readiness assessments before implementation.
- Recognize that infrastructure capabilities and user perceptions both critically influence adoption.
- Invest in change management and attitude-building alongside technical infrastructure- Consider local, political, cultural, and regulatory contexts when adapting our recommendations- Pilot projects in limited scope before full-scale deployment to build confidence and demonstrate value.

Future Research Directions

Future research should address current limitations by:

- Conducting longitudinal studies tracking Tehran Governorate (or similar organizations) through actual implementation to examine how perceptions evolve and predict the actual usage.
- Expanding to multi-organizational comparative studies across different

governorates to assess the model generalizability.

- Implementing qualitative methods (interviews, focus groups) to deeply understand organizational dynamics, political considerations, and cultural factors influencing the adoption.
- Extending the theoretical model to include context-specific variables such as political support, bureaucratic processes, inter-agency collaboration, and citizen participation levels.
- Examining the front-line staff perspectives alongside the managerial views.
- Investigating the moderating effects of organizational size, leadership style, technological maturity, and governance structures.
- Exploring the citizen perspectives on government use of big data analytics and privacy concerns.
- Conducting comparative studies across different national and cultural contexts.

In conclusion, this research demonstrated that Tehran Province Governorate's readiness to adopt big data analytics for CRM depends on both technological infrastructure and user acceptance factors. Successful implementation requires comprehensive attention to infrastructure development, human capacity building, organizational culture change, and governance frameworks. While findings are context-specific, the methodological approach and theoretical framework may guide similar assessments in other governmental organizations.

REFERENCE

Al Batayneh, R. M., Taleb, N., Said, R. A., Alshurideh, M. T., Ghazal, T. M., & Alzoubi, H. M. (2021, May). IT governance framework and smart services integration for future development of Dubai infrastructure utilizing AI and big data, its reflection on the citizens standard of living. In *Proceedings of the International Conference on Artificial Intelligence and Computer Vision (AICV2021)* (pp. 235-247). Springer International Publishing.

Al-Ateeq, B., Sawan, N., Al-Hajaya, K., Altarawneh, M., & Al-Makhadmeh, A. (2022). Big data analytics in auditing and the consequences for audit quality: A study using the technology acceptance model (TAM). *Corporate Governance and Organizational Behavior Review, 6*(1), 64-78. https://doi.org/10.22495/cgobrv6i1p6.

Aldholay, A., Isaac, O., Jalal, A. N., Anor, F. A., & Mutahar, A. M. (2022). Factors that accelerate the rise of acceptance of big data platforms for academic teaching: Personal innovativeness as moderating variable. In *Proceedings of International Conference on Emerging Technologies and Intelligent Systems* (pp. 227-243). Springer International Publishing.

Al-Dossari, S., Mokhtar, U. A., & Abdul Ghani, A. T. (2023). Factor influencing the adoption of big data analytics: A systematic literature and experts review. *SAGE Open, 13*(4), 21582440231217902. https://doi.org/10.1177/21582440231217902.

Allen, B., Tamindael, L. E., Bickerton, S. H., & Cho, W. (2020). Does citizen coproduction lead to better urban services in smart cities projects? An empirical study on e-participation in a mobile big data platform. *Government Information Quarterly, 37*(1), 101412. https://doi.org/10.1016/j.giq.2019.101412.

Alyusuf, I. Y., & Al-Rahmi, W. M. (2022). Big data analytics adoption via lenses of Technology Acceptance Model: Empirical study of higher education. *Entrepreneurship and*

Sustainability Issues, 9(3), 399-417. https://doi.org/10.9770/jesi.2022.9.3(25).

Anshari, M., Almunawar, M. N., Lim, S. A., & Al-Mudimigh, A. (2018). *Applied computing and informatics*.

Anshari, M., Almunawar, M. N., Lim, S. A., & Al-Mudimigh, A. (2019). Customer relationship management and big data enabled: Personalization & customization of services. *Applied Computing and Informatics, 15*(2), 94-101. https://doi.org/10.1016/j.aci.2018.12.001.

Arghashi, V., & Yuksel, C. A. (2022). Interactivity, inspiration, and perceived usefulness! How retailers' AR-apps improve consumer engagement through flow. *Journal of Retailing and Consumer Services, 64*, 102756. https://doi.org/10.1016/j.jretconser.2021.102756.

Asadpoor, E., Mohammadi, F., & Rahimi, A. (2024). Acceptance of artificial intelligence among medical students in Iran: A Technology Acceptance Model approach. *Medical Education Journal, 18*(2), 45-58.

Ashiku, L., Al-Amin, M., Madria, S., & Dagli, C. (2021). Machine learning models and big data tools for evaluating kidney acceptance. *Procedia Computer Science, 185*, 177-184.

Aworh, M. K., Kwaga, J. K. P., & Okolocha, E. C. (2021). Assessing knowledge, attitude, and practices of veterinarians towards antimicrobial use and stewardship as drivers of inappropriate use in Abuja, Nigeria. *One Health Outlook, 3*(1), 25. https://doi.org/10.1186/s42522-021-00058-3.

Barham, H., & Daim, T. (2020). The use of readiness assessment for big data projects. *Sustainable Cities and Society, 60*, 102233. https://doi.org/10.1016/j.scs.2020.102233

Baycan, T., & Yigitcanlar, T. (2024). Smart city governance: Assessing modes of active citizen engagement. *Regional Studies, 59*(1), 2399262. https://doi.org/10.1080/00343404.2024.2399262.

Beesmart City. (2024, February 8). *How smart cities are boosting citizen engagement*. https://www.beesmart.city/en/smart-city-blog/how-smart-cities-boost-citizen-engagement.

Beesmart City. (2025, January 15). *5 citizen engagement solutions that smart cities should know*. https://www.beesmart.city/en/smart-city-blog/top-civic-engagement-solutions-smart-cities.

Cabrera-Sánchez, J.-P., & Villarejo-Ramos, Á. F. (2020). Acceptance and use of big data techniques in services companies. *Journal of Retailing and Consumer Services, 52*, 101888. https://doi.org/10.1016/j.jretconser.2019.101888.

Caffaro, F., Micheletti Cremasco, M., Roccato, M., & Cavallo, E. (2020). Drivers of farmers' intention to adopt technological innovations in Italy: The role of information sources, perceived usefulness, and perceived ease of use. *Journal of Rural Studies, 76*, 264-271. https://doi.org/10.1016/j.jrurstud.2020.04.028.

Cao, Y., & Kang, M. (2025). Digital dialogue in smart cities: Evidence from public concerns, government responsiveness, and citizen satisfaction in China. *Cities, 158*, 105174. https://doi.org/10.1016/j.cities.2025.105174.

Chamoso, P., González-Briones, A., De La Prieta, F., Venyagamoorthy, G. K., & Corchado, J. M. (2020). Smart city as a distributed platform: Toward a system for citizen-oriented management. *Computer Communications, 152*, 323-332. https://doi.org/10.1016/j.comcom.2020.01.059.

Civita App. (2025, February 18). *Why every city needs a Citizen Relationship Management mobile app*. https://civitaapp.com/citizen-relationship-management-mobile-app/.

Costa, R., Fernandes, G., & Silva, M. (2025). Understanding recruiters' acceptance of artificial intelligence: Insights from the Technology Acceptance Model. *Applied Sciences, 15*(2), 746. https://doi.org/10.3390/app15020746.

Fahmfam Ghodsiyeh and Hamidi Hodjat Fahm Fāmm, & Hamidi. (2018). Factors influencing the development and management of smart cities using an integrated approach of big data technologies, Internet of Things, and cloud computing. *Journal of Information Processing and Management, 34*(2), 557-584. [in Persian]

Fallahi Modaresi, S., & Zarei, A. (2022). A study on the factors influencing the adoption of big data technology in the tourism industry using the Technology Acceptance TOE framework (case study: Tourism industry jobs in Shiraz). *Journal of Tourism and Development*. [in Persian]

Faridoon, L., Liu, W., & Spence, C. (2024). The impact of big data analytics on decision-making within the government sector. *Big Data, 3*(X), 1-17. https://doi.org/10.1089/big.2023.0019.

Galvez Rachel. (2024, July 12). *Data governance adoption has risen dramatically - Here's how*. https://www.precisely.com/data-integrity/2025-planning-insights-data-governance-adoption-has-risen-dramatically/.

Ghali, E. A. A., Dominic, P. D. D., Fati, S. M., Muneer, A., & Ali, R. F. (2021). The assessment of big data adoption readiness with a technology–organization–environment framework: A perspective towards healthcare employees. *Sustainability, 13*(15), 8379. https://doi.org/10.3390/su13158379.

Hamta, N., Mohammadzadeh, Y., Hemati, M., & Dehghanzadeh, R. (2020). A study of user attitudes toward the medical imaging storage and transfer system in the educational and medical centers of Qom based on the Technology Acceptance Model (TAM). *Qom University of Medical Sciences Journal, 14*(6), 1-8. [in Persian]

Hossin, M. A., Du, J., Mu, L., & Asante, I. O. (2023). Big data-driven public policy decisions: Transformation toward smart governance. *SAGE Open, 13*(4), 21582440231215123. https://doi.org/10.1177/21582440231215123.

Ibrahim, F., Münscher, J.-C., Daseking, M., & Telle, N.-T. (2024). The technology acceptance model and adopter type analysis in the context of artificial intelligence. *Frontiers in Artificial Intelligence, 7*, 1496518. https://doi.org/10.3389/frai.2024.1496518.

Infosys Public Services. (2024). *Big data analytics for government sector*. https://www.infosyspublicservices.com/insights/blogs/bytes-better-policies.html.

Iriani, S. S., & Andjarwati, A. L. (2020). Analysis of perceived usefulness, perceived ease of use, and perceived risk toward online shopping in the era of Covid-19 pandemic. *Systematic Reviews in Pharmacy, 11*(12), 313-320.

Islami, M. M., Asdar, M., & Baumassepe, A. N. (2021). Analysis of perceived usefulness and perceived ease of use to the actual system usage through attitude using online guidance application. *Hasanuddin Journal of Business Strategy, 3*(1), 52-64.

Ju, J., Liu, L., & Feng, Y. (2018). Citizen-centered big data analysis-driven governance intelligence framework for smart cities. *Telecommunications Policy, 42*(10), 881-896. https://doi.org/10.1016/j.telpol.2018.01.003.

Kar, A. K., & Dwivedi, Y. K. (2020). Theory building with big data-driven research -- Moving away from the "What" towards the "Why". *International Journal of Information Management, 54*, 102205. https://doi.org/10.1016/j.ijinfomgt.2020.102205.

Kasilingam, D. L. (2020). Understanding the attitude and intention to use smartphone chatbots for shopping. *Technology in Society, 62*, 101280. https://doi.org/10.1016/j.techsoc.2020.101280.

Kumar, V., & Reinartz, W. (2018). *Customer relationship management*. Springer International Publishing.

Lampropoulos, G., Siakas, K., Viana, J., & Reinhold, O. (2022). Artificial intelligence, blockchain, big data analytics, machine learning and data mining in traditional CRM and social CRM: A critical review. In *2022 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)* (pp. 504-510). IEEE.

Li, C., Chen, Y., & Shang, Y. (2022). A review of industrial big data for decision making in intelligent manufacturing. *Engineering Science and Technology, an International Journal, 29*, 101021. https://doi.org/10.1016/j.jestch.2021.06.001.

Maroufkhani, P., Tseng, M.-L., Iranmanesh, M., Ismail, W. K. W., & Khalid, H. (2020). Big data analytics adoption: Determinants and performances among small to medium-sized enterprises. *International Journal of Information Management, 54*, 102190. https://doi.org/10.1016/j.ijinfomgt.2020.102190.

Mikalef, P., Krogstie, J., Pappas, I. O., & Pavlou, P. (2020). Exploring the relationship between big data analytics capability and competitive performance: The mediating roles of dynamic and operational capabilities. *Information & Management, 57*(2), 103169. https://doi.org/10.1016/j.im.2019.103169.

Na, S., Heo, S., Han, S., Shin, Y., & Roh, Y. (2022). Acceptance model of artificial intelligence (AI)-based technologies in construction firms: Applying the Technology Acceptance Model (TAM) in combination with the Technology–Organisation–Environment (TOE) framework. *Buildings, 12*(2), 90. https://doi.org/10.3390/buildings12020090.

Naeem, M., et al. (2022). Trends and future perspective challenges in big data. In *Advances in intelligent data analysis and applications* (pp. 309-325). Springer Singapore.

Naeem, M., Jamal, T., Diaz-Martinez, J., Butt, S. A., Montesano, N., Tariq, M. I., ... & De-La-Hoz-Valdiris, E. (2022, November). Trends and future perspective challenges in big data. In *Advances in intelligent data analysis and applications: Proceeding of the sixth euro-China conference on intelligent data analysis and applications, 15–18 October 2019, Arad, Romania* (pp. 309-325). Singapore: Springer Singapore.

Noori, A., & Emamvirdi, G. (2015). Designing a model to assess readiness for digital/electronic readiness in Iranian service organizations using Data-Driven Theory Construction (DDD) method. *Decision Engineering, 3*(1), 61-86. [in Persian]

Noori, A., Hatami, Z., & Ebrahimiān, H. (2017). Factors influencing the adoption of information technology and its impact on human resources. *Human Resource Management Research, 9*(4), 127-153. [in Persian]

Oubdi, L., & El-Mekkaoui, O. (2026). The adoption of AI tools in doctoral studies: An extended TAM framework. In S. D. Carter, A. Smith-Hunter, & L. Best (Eds.), *Impact of artificial intelligence (AI) and the global financial crisis on development in Africa* (pp. 47-63). Springer International Publishing. https://doi.org/10.1007/978-3-031-94518-2_4.

Pangarkar Tajammul. (2024, August 13). *Data governance statistics and facts (2025): Emerging technologies, challenges and adoption, AI, ROI, and data quality insights*. https://electroiq.com/stats/data-governance/.

QodeaCTS. (2024). *Data analytics in the public sector: Using data to improve services*. https://qodea.com/resources/data-analytics-in-the-public-sector-using-data-to-improve-services/.

Rahman, M. N. (2020). *Exploring the factors influencing big data technology acceptance* [Doctoral dissertation, Portland State University]. PDXScholar.

Randi Vahid, Khoon., Siyavosh Mohsen, & Bamsoomi Behrooz. (2014). Factors influencing online customer purchase behavior in Iran with respect to the Technology Acceptance Model (TAM). *Management of Development and Transformation*, 109-118. [in Persian]

Rattletech. (2021, October 14). *How a Citizen Relationship Management solution can transform your city*. https://www.rattletech.com/citizen-relationship-management-solution-transform-city/

Sarker, I. H. (2021). Data science and analytics: An overview from data-driven smart computing, decision-making and applications perspective. *SN Computer Science, 2*(5), 377. https://doi.org/10.1007/s42979-021-00765-8.

Shukla, D. (2024, January 23). Big but not scary: How to use big data to shape government policy and delivery. *Global Government Forum*. https://www.globalgovernmentforum.com/big-but-not-scary-how-to-use-big-data-to-shape-government-policy-and-delivery/.

Siagian, H., Tarigan, Z. J. H., Basana, S. R., & Basuki, R. (2022). *The effect of perceived security, perceived ease of use, and perceived usefulness on consumer behavioral intention through trust in digital payment platform*. Petra Christian University.

Siagian, H., Tarigan, Z. J. H., Basana, S. R., & Basuki, R. (2022). The effect of perceived security, perceived ease of use, and perceived usefulness on consumer behavioral intention through trust in digital payment platform. *International Journal of Data & Network Science*, *6*(3), 861–874.

Singh, P. D. (2024). Generative AI through the lens of Technology Acceptance Model. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4953174.

Smith, A. M. (2024, September 15). Data governance trends in 2025. *DATAVERSITY*. https://www.dataversity.net/articles/data-governance-trends-in-2025/.

Xu, Z., Zhang, J., Zhang, Z., Li, C., & Wang, K. (2020). How to perceive the impacts of land supply on urban management efficiency: Evidence from China's 315 cities. *Habitat International, 98*, 102145. https://doi.org/10.1016/j.habitatint.2020.102145.

Journal of
Knowledge Economy Studies

https://kes.hmu.ac.ir

Hazrat-e Masoumeh University

Ranking the Challenges of Cryptocurrency Development in Iran

Article Type:
Research Article

Mohammad Ghaffary Fard[1*]
Ahlul Bayt international university
Corresponding Author
E-mail: ghaffary2@yahoo.com

Seyed Ali Akbar Sadat[2]
Ahlul Bayt international university
E-mail: sadat1999.1378@gmail.com

Mohsen Jafarian[3]
Faculty of Business and Management, UCSI University, Kuala Lumpur
E-mail: mohsenjafarian@gmail.com

Received: 06 October 2025
Revised: 07 November 2025
Accepted: 26 December 2025
Available Online: 27 December 2025

Abstract

The emergence of cryptocurrencies in recent years presents a novel phenomenon in digital economics, offering both opportunities and challenges for nations. The main aim of this study was to examine the challenges of cryptocurrency development, with a specific focus on the contents of the Popular Government's Transformation Document. This research is applied in purpose and descriptive in nature, utilizing a field-based questionnaire for data collection. In the initial phase, challenges were identified through conducting library research, specifically leveraging the Transformation Document. Subsequently, these challenges were evaluated by experts using the Analytic Hierarchy Process (AHP), and ultimately ranked with the aid of Expert Choice software. The findings, validated by a strong inconsistency rate of 0.06, revealed that the weakness in macro-management of cryptocurrencies holds the first rank among the main challenges, with a coefficient of 0.429. Conversely, the increasing share of hidden mining in the country's cryptocurrency production market ranks last, with a coefficient of 0.114. Given these results, policymakers must prioritize the immediate establishment of a unified and powerful command structure, coupled with a clear clarification of responsibilities. Simultaneously, core strategies must integrate the protection of small capital and safeguarding public trust in the financial system to effectively mitigate social risks and ensure financial stability.

Keywords

Cryptocurrency, Ranking, Popular government's transformation document, Analytical hierarchy process.

Cite this article: Ghaffary Fard, M., Sadat, A. A, & Jafarian, M., (2026). Ranking the challenges of cryptocurrency development in iran. *Journal of Knowledge Economy Studies (JKES)*, 3(1), 121-139. DOI: http//doi.org/10.22034/kes.2026.2073895.1086

Online ISSN: 3060-7329

1. Introduction

Money is considered as one of the fundamental pillars of human life, with a history spanning thousands of years, utilized in various forms throughout different eras. With the advancement of new technologies and the expansion of the digital realm, a new type of asset called "cryptocurrency" or "digital currency" has emerged, which is increasingly growing. Cryptocurrencies are a form of digital asset that operates based on Blockchain technology. This technology is a decentralized and distributed platform that, ensures the security of transactions and controls the generation of new monetary units by leveraging advanced cryptographic algorithms (Varmaziari et al., 2024). Among the most important features of cryptocurrencies are their unforgeability, independence from geographical borders, transparency, high speed, and low transaction costs (Saeidi et al., 2024).

The first known cryptocurrency is Bitcoin, introduced in 2008 by an anonymous individual or group using the pseudonym "Satoshi Nakamoto." As the first digital currency, Bitcoin paved the way for the development of thousands of other cryptocurrencies, which are known as "Altcoins." The term Altcoin is derived from the combination of "Alternative" and "Coin," meaning an alternative coin or cryptocurrency (Hashem Nejad et al., 2024).

Cryptocurrencies, with their unique nature and characteristics, are distinguished from other forms of digital money and traditional fiat currencies, and possess specific strengths and weaknesses. Now, more than a decade after the emergence of the first digital currency, we witnessed that the number of these cryptocurrencies has exceeded ten thousand, indicating their increasing popularity and value in international markets. Consequently, familiarity with the concept of cryptocurrencies seems widespread today. Therefore, it is essential for governments, to optimally leverage the benefits of this nascent phenomenon by formulating and implementing comprehensive policies and Laws while precisely identifying its shortcomings to prevent potential damages to the national economy (Zare, 2022) .

Despite conducting extensive research on cryptocurrency challenges—such as legal ambiguities, security risks (e.g., money laundering, price volatility), high energy consumption, and opportunities like sanction circumvention—Iran's policymaking system lacks a prioritized analytical framework aligned with its unique governance and economic conditions, including sanctions and the need for economic stability. The People's Government Transformation Document, the guiding framework of Iran's 13th administration, prioritizes justice, employment, economic stability, and anti-corruption.

In fulfillment of the promise made at the beginning of the formation of the 13th government and in line with the realization of the goals set forth in the Constitution, the Vision Document of the Islamic Republic of Iran in the Horizon 1404, the general policies of the system and meeting the needs of the people, resolving the fundamental issues, the "People's Government Transformation Document" became the government's program and policy. It became the basis for the action of the executive branch,ministers, and executive agencies. The People's Government Transformation Document was developed in accordance with the current situation, achieving the desired situation. It was developed in line with the general policies of the system in the form of priority basic issues with the cooperation of ministers, vice presidents, and officials of executive agencies. Moreover, the collective efforts of experts from the field and universities, think tanks, and research institutes were essential in guidinge and creatinge coherence in the directions of the programs of the various executive departments of the government. The responsibility for implementing the provisions of the Transformation Document in relevant cases lies with the ministers and the highest level officials in the relevant agency . They are obliged to be accountable (Iran, 2022) .

This study aimed at identifyinge the cryptocurrency challenges which most threaten these goals. The existing studies described these challenges but lack quantitative rankings based on the Document's criteria. This study used the Analytical Hierarchy Process (AHP) and Expert Choice software to rank and weight these challenges, aiming to guide the effective policy prioritization.

Therefore, the present article examined the challenges hindering the development of cryptocurrencies based on the Popular Government's Transformation Document. In this study, the Analytical Hierarchy Process (AHP) method has been used for the decomposition, analysis, and comparison of challenges. Finally, these challenges have been ranked and weighted by utilizing Expert Choice software. The central question of this research study is as follows:

Q. Which of the challenges in cryptocurrency development are of greater importance and are identified as the most significant challenges, from the perspective of experts and specialists in this field?

2. Theoretical Foundations

With the expansion of new technologies in the field of economy and financial services, the concept of cryptocurrencies, as one of the new pillars of the digital economy, has found a special place in economic studies and policymaking. The current research study aimed to prioritize and rank the challenges of cryptocurrency development based on the "Popular Government Transformation Document". This document identifies four main challenges for cryptocurrency development, each of which has been defined and examined in the theoretical foundations of this study.

2.1 Cryptocurrency

Cryptocurrencies, often known as digital currencies, are a collection of digital assets that utilize cryptography to secure their financial transactions. These currencies have created a new online payment system that offers innovative features and capabilities. Unlike traditional financial systems, these systems are not connected to a centralized financial intermediary or institution, providing peer-to-peer online payment capabilities(Li & Huang, 2019). These distributed systems do not require the central, supervisory, and physical control present in conventional financial systems (Corbet et al., 2019). A digital currency or cryptocurrency is a type of money that facilitates exchanges between users using cryptographic algorithms. This cryptography ensures the security of digital money and prevents its fraud and counterfeiting (like physical banknotes such as dollar). Many digital currencies are based on a decentralized network built on Blockchain technology (a distributed ledger managed by a virtual network). Unlike fiat currencies, which are printed by the government, digital currencies do not have private ownership. This feature makes them immune to manipulation and intervention by entities and organizations. Digital currencies are essentially encrypted currencies designed with encrypted protocols, aiming to reduce fraud and prevent currency counterfeiting and scams. The most important feature of digital currency is its decentralized nature, meaning no specific entity or organization supervises and controls it (Zulhuda, 2017). In 2008 an individual under the pseudonym Satoshi Nakamoto introduced a distributed digital financial asset called Bitcoin for the first time to the world. In October 2009, the initial price of each Bitcoin was set at $0.000764 (Ammous, 2018). Bitcoin, as the first cryptocurrency, is currently recognized as the leading and most successful decentralized digital money(Babazadeh et al., 2021).

2.2 Popular Government Transformation Document

The Popular Government Transformation Document is a codified plan of the 13th government of the Islamic Republic of Iran for achieving the desired state in the country. This document has been compiled in line with the objectives of the Constitution, the 2025 Vision Document, and the general policies of the system, leveraging the statements of the Supreme Leader of the Islamic Revolution. The main goals of this document include justice-orientation and anti-corruption, knowledge and scientific resurgence, public participation with youth leadership, and family-centricity. The main axes of this document are structured into nine chapters including: production and employment, investment and financial system, public finance system, infrastructure, social affairs and health, education, culture and art, administrative and legal system, foreign policy, and security. This document serves as the basis for the executive branches, ministers, and executive agencies,.The responsibility for implementing its provisions lies with the highest authority in each agency (Iran, 2022) .

2.3 Challenges of Cryptocurrency Development in the Popular Government Transformation Document

2.3.1 Weakness in Macro-Management of Cryptocurrencies

The absence of clear and comprehensive regulations for cryptocurrencies at the macro level is considered as one of their main management problems. This has led to different countries adopting varying approaches, from complete prohibition to limited acceptance. Interestingly, the global perspective on this digital phenomenon is diverse; some countries consider cryptocurrencies as a serious threat and have declared their use illegal, while other countries deem them beneficial and have imposed tax laws on them. Most developed countries have moved towards drafting laws in this area with a constructive (rather than prohibitive) approach. Another important challenge arising from cryptocurrencies is their multifaceted nature, which causes ambiguity in understanding their essence. The multifaceted nature of cryptocurrencies leads to their connection with most economic and non-economic organizations in the country. This thematic dependency, in the absence of a centralized body, creates a serious obstacle to the management and regulation of cryptocurrencies (Rostami et al., 2024).

2.3.2 High Investment Risk for the Public in the Cryptocurrency Market

The importance of economic security in the 21st century has taken on a broader and more complex meaning. In the previous two centuries, military and political security held the most importance, but in the current century, economic security has taken precedence. Today, special importance is given to the economic security index in evaluating the development level of countries (Ahmadpour, 2024).

Bitcoin is considered a financial and electronic innovation that has expanded and gained popularity in the past five years. Despite its boom, Bitcoin still faces challenges such as extremely high volatility in its value, non-recognition by the Central Bank, absence of codified laws regarding Bitcoin and legal silence, security issues, threats to the real economy, and impacts on the economy and traditional money supply, which pose serious concerns for investors in this field (Bagban, 2019).

2.3.3 Increasing the Share of Hidden Mining in the Total Cryptocurrency Production Market in the Country

Cryptocurrency mining in Iran has become a political issue linked to the stability of the country's electricity grid. Also, the biggest risk in this area has been illegal mining and legal action against it. So far, most discoveries in Bitcoin mining have involved miners operating without a license. From the beginning of the crackdown on illegal miners between 2019 and 2021, a total of 221,163 illegal cryptocurrency mining devices were identified and seized, with a power consumption equivalent to 621 megawatts. However, not all these devices were simultaneously active in the electricity grid. The country's most explicit action in cryptocurrency mining policy was the imposition of a mining ban in the summer of 2021. However, this policy is practically considered as a failure because the absolute monthly terahash of Iran in the Bitcoin network only decreased from 6.94 million terahashes on May 1, 2021, to 3.75 million terahashes on August 1, 2021. In other words, half of the cryptocurrency miners were secretly engaged in mining despite the government's order. The power consumption of these miners was, at worst, less than 800 megawatt-hours, while the country's electricity grid deficit was over 10,000 megawatt-hours. At worst, less than 10% of the country's electricity deficit can be attributed to cryptocurrency mining (Rajabi & Saberi, 2022).

2.3.4 Lost Opportunities in Utilizing the Strategic Capacities of Cryptocurrencies in Domestic and International Payments and Exchanges

In recent years, international sanctions have led to restrictions on financial and banking exchanges in the international financial system. This has resulted in the non-repatriation of revenues into the country and budget provision, which is one of the country's most important priorities in international trade. Using cryptocurrencies can serve as a suitable platform to bypass financial and banking sanctions. In a research study conducted by Babazadeh et al., (2021), six groups of key indicators were identified and prioritized, which include developing laws and regulations related to cryptocurrencies, creating the necessary software and hardware platforms, designing and offering national cryptocurrencies, promoting and expanding the use of cryptocurrencies, and supporting the cryptocurrency mining processes. These can serve as an efficient guide for the Iranian government and international economic actors affected by sanctions, in order to counter financial and banking restrictions (Babazadeh et al., 2021).

3. Background of the Research

3.1 Governance and Regulatory Frameworks

The absence of consistent and comprehensive legal and regulatory frameworks represents a significant barrier to cryptocurrency adoption globally and domestically. Internationally, inconsistent tax policies across jurisdictions create confusion and hinder compliance (Adhikari et al., 2025). Similarly, the rapid pace of technological innovations in cryptocurrencies outpaces regulatory responses, enabling illicit activities such as money laundering and terrorist financing due to inconsistent global anti-money laundering (AML) regulations(He et al., 2024; Oye et al., 2025). Domestically, legal ambiguities complicate efforts to address crimes like fraud and forgery in cryptocurrency transactions (Sadeghi et al., 2024); (Shamsi et al., 2024). From a jurisprudential perspective, some scholars argue that cryptocurrency transactions involve *gharar* (deception and uncertainty) (HabibianNaqibi et al., 2020), while others asserted their permissibility within religious frameworks (Taj Langerudi & Dehdar, 2024). Both international and domestic studies emphasized the urgent need for having robust regulatory frameworks. Domestically, a shift from prohibitive policies to regulated management is recommended to align with national policy objectives(Aref et al., 2024) ;(babak et al., 2024); (Ghaemi Asl et al., 2024).

The integration of Internet of Things (IoT) systems within food supply networks encounters substantial obstacles that limit its capacity to enhance operational transparency, productivity, and ecological balance. This research delineated and examined these impediments through conducting an in-depth investigation of Kaleh Company, a prominent food producer in Iran, applying the DEMATEL analytical framework. The results indicated that effective IoT deployment necessitates integrated approaches that concurrently tackle diverse areas, such as establishing universal protocols, creating economic incentives, and enhancing employee competencies (Fathi et al., 2025).

The present study investigated the obstacles impacting the implementation of Artificial Intelligence within the desalination supply chain. This sector is recognized as a fundamental solution for addressing global water scarcity. Based on a systematic literature review and consultations with experts, sixteen primary barriers were identified and subsequently analyzed using the MICMAC method. The findings from this analysis revealed four factors that function as critical barriers and primary bottlenecks to the successful deployment of AI. These factors include insufficient funding and capital, absence of standardization and system interoperability, a deficit of specialized skills and qualified personnel, and concerns regarding data security and privacy (Sadeghi et al., 2025)

2.3 Energy, Mining, and Infrastructure Challenges

The energy-intensive nature of cryptocurrency mining, particularly through proof-of-work (PoW) consensus protocols, poses significant environmental and infrastructural challenges. Globally, the high energy consumption of PoW raises environmental concerns (Stoll et al., 2019). Proposed solutions, such as proof-of-stake (PoS), aim to reduce energy use, though scalability and adoption debates persist (Zimba et al., 2025). Domestically, scalability and technical complexity are identified as major barriers to Blockchain infrastructure development (Mohammadi Fateh & Salarnejad, 2022). Innovative solutions, such as integrating cryptocurrency mining with power plants to attract private investment in the electricity sector, have been proposed to address domestic energy challenges (Larijani & Taheri, 2022). Both global and domestic research underscored the need for technological advancements and infrastructure investments to ensure sustainable cryptocurrency development.

3.3 Financial Risks and Economic Stability

Cryptocurrencies introduce significant financial risks that undermine the investor's confidence and economic stability. Internationally, high-profile fraud cases, such as the FTX[1] scandal, highlight the tangible threat of fraud to investors (Kerr et al., 2023), with negative experiences further eroding trust (Lourie et al., 2023). Additionally, the substitution of cryptocurrencies for bank deposits could destabilize banking systems and heighten the risks of financial crisis (Chen & Phelan, 2025);(Zheng, 2025). Domestically, the anonymity and decentralization of cryptocurrencies exacerbate economic crimes, including laundering the money and expanding the shadow economy (Kashian & Parnian, 2021);(Rostami et al., 2024). Despite these risks, opportunities exist, particularly domestically, where cryptocurrencies could serve as a strategic tool to circumvent banking and financial sanctions (Babazadeh et al., 2021);(Zare, 2022). Both global and domestic studies advocate for balanced regulatory measures to mitigate the risks while capitalizing on economic opportunities.

The Research Gap and the Study's Contribution

While the reviewed studies provided a comprehensive analysis of cryptocurrency challenges and opportunities, a critical gap persists: none explicitly aligned their findings with the "People's Government Transformation Document", a pivotal national policy framework guiding objectives such as justice, employment, economic stability, and anti-corruption. This study addresses this gap by systematically evaluating and prioritizing cryptocurrency challenges using the Expert Choice software. By integrating global and domestic perspectives thematically, this research offers precise and, evidence-based recommendations to foster cryptocurrency development within a robust governance framework, ensuring alignment with national policy goals.

4. Methodology

4.1 Design

The present research aimed to identify and prioritize the main challenges of cryptocurrencies, by identifying and ranking the challenges of cryptocurrency

1. Futures Exchange

development in Iran. The study is applied in purpose and descriptive-empirical in nature. It used a survey-based method with a questionnaire. In the initial phase of the study, the decision-making hierarchy was established using library studies and explicit content analysis. In this academic research, the Popular Government Transformation Document Sanade Tahavvol Dolat Mardomi has been centrally chosen as the core analytical framework, being a macro-strategic document grounded in three essential pillars of achieving developmental justice through equitable distribution of economic opportunities, structural empowerment of low-income strata within the Resistance Economy Eghtesad Moghavemati framework, and the systematic application of modern technologies for enhancing transparency and developing the financial system. This framework's selection is fundamentally driven by three reasons. First, the document's explicit emphasis in Section 4-2 on the mandatory organization of new financial institutions under the Resistance Economy; second, the stipulation in Section 5-1 to establish criteria for prioritizing obstacles to technology development based on justice-oriented and popular development indicators; and third, its provision of an indigenous framework aligned with Iran's upstream development documents, which facilitates converting abstract challenges in cryptocurrency development into practical and actionable policy solutions, thereby ensuring both the research's theoretical coherence and the enforceability and practical application of its findings within the national policy-making system (Iran, 2022). The "Transformation Document of the People's Government", focusing on its specialized sections related to cryptocurrencies, has clearly outlined the key challenges and their sub-challenges, shifting the governance approach from a passive state to an active, opportunity-oriented one with risk management. This analysis ultimately led to the identification of four main criteria and nine sub-criteria, which directly reflect the objectives and threats stated in the document. The primary data collection tool was a pairwise comparison questionnaire, systematically designed based on the four main criteria and nine sub-criteria extracted from the Transformation Document. This questionnaire contained twelve sets of pairwise comparisons where, in each section, one challenge was compared against another. The preference of one challenge over another was measured using a relative numerical scale relative to the ultimate goal.

4.2 Sampling and Expert Validation

In the data collection stage, to ensure the scientific and specialized competence of the sample, a purposive non-random sampling method was employed, allowing only specialists with profound theoretical knowledge and sufficient practical experience in the research domain to enter the analysis process. This Expert Panel comprised 13 outstanding specialists, all of whom possessed a minimum of five years of professional experience in related fields. Academically, the group included 3 individuals holding a Ph.D. degree (in Economics and Management disciplines) and 10 individuals with a Master's degree (in Information Technology Management, Financial Engineering, and Business Management). The specialized composition of the panel was meticulously balanced, consisting of 3 senior experts from the Central Bank of the Islamic Republic of Iran (as the principal regulatory authority), 7 managers and executive activists from the digital currency industry and cryptocurrency exchanges, and 3 university faculty members (focused on macro-economic analysis). This specialized structure not only furnishes the highest level of knowledge and credibility for the qualitative judgments within the Analytic Hierarchy Process (AHP) model, but also guarantees the direct alignment of the results with Iran's executive realities and regulatory environment, thereby significantly boosting the enforceability of the research findings.

4.3 AHP Procedure

The data obtained from the expert survey were analyzed using the Analytical Hierarchy Process method. This method, considered one of the advanced approaches in decision science, converts the qualitative and subjective judgments of experts into quantifiable and measurable values and is used to determine relative weights and prioritize complex hierarchical structures. All calculations and analyses were performed using the specialized Expert Choice software. This research framework, relying on the robustness of the overarching document and the methodological precision of the Analytical Hierarchy Process, yields valid and actionable outputs for the country's policy-making system. By providing targeted quantitative weighting of the challenges, it addresses the existing gap in the research literature of this field. The challenges facing cryptocurrency development, which include four main criteria and nine sub-criteria, are presented in the table below.

Table 1.
The Main Challenges and Factors

Row	Main challenges	Symbol	Row	Factors	Symbol
1	Weakness in macro-level management of cryptocurrencies	A	1	Inefficient governance structure in actively confronting cryptocurrencies; distributed ledger technology offices	a1
			2	Weakness in monitoring, controlling, supervising,, and self- regulating the market for producing, storing, and exchanging cryptocurrencies	a2
2	High risk to people's investment in the cryptocurrency market	B	1	Motivation and excitement of people to escape inflation and purchase non-Rial assets with high liquidity	b1
			2	Inadequate regulation in the area of asset management policies and custody of cryptocurrencies	b2
			3	Weakness of policies supporting domestic production and export within the cryptocurrency ecosystem	b3
3	Rising share of hidden/mined extraction in the total cryptocurrency production market in the country	C	1	Loss of an intelligent monitoring and control structure for electricity consumption patterns	c1
			2	Inappropriate pricing of energy carriers and discrimination in dealing with the cryptocurrency mining industry compared to other industries	c2
4	Loss of opportunities in leveraging the strategic capacities of cryptocurrencies in domestic and international payment and exchange domains	D	1	Liberation and ambiguity/uncertainty of cryptocurrency exchange and payment infrastructures domestically	d1
			2	Excessive imports and capital outflow, and non-activation of exports	d2

(Source: The Researcher's Findings)

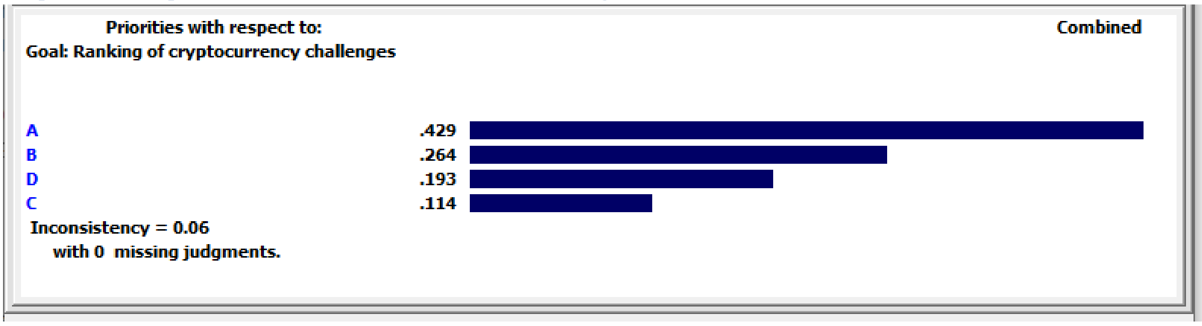
5. Data Analysis and Findings

The study included four primary criteria and nine subsidiary criteria. To weight these criteria, the Analytic Hierarchy Process (AHP) was employed. For this purpose, a questionnaire consisting of main and sub-criteria was distributed among 13 experts. The experts used a 1-9 Likert scale to determine the relative importance of the criteria. Finally, using the Expert Choice software, the weights of the main and sub-criteria were calculated and obtained.

5.1 The Main Challenges in Cryptocurrency Development

The results of AHP conducted using the Expert Choice software, based on input from 13 domain experts, provided a strategic ranking of challenges in the cryptocurrency sector. In addition to determining the weight and priority of each challenge, the methodological reliability of the analysis was confirmed by an inconsistency rate of 0.06—below the standard threshold of 0.1. This statistical indicator demonstrated that the pairwise comparisons made by the experts were logically consistent, allowing full confidence in the study's results. According to the findings, weakness in macro-level cryptocurrency management, with a weight of 0.429, ranked first. This clearly indicates that the primary threat in this domain lies not in the financial or infrastructural nature of cryptocurrencies, but in governance gaps and inefficient decision-making structures. This structural inefficiency directly contradicts the proactive and opportunity-driven management approach emphasized in the Transformation Document and is identified as the root cause of other challenges. In second place, with a weight of 0.264, was the high risk of public investment. This reflects the priority of safeguarding economic stability and public rights. From the experts' perspective, the unregulated influx of small-scale public capital into high-risk markets poses a serious socioeconomic threat that challenges public trust in the financial system. The third challenge, with a weight of 0.193, was the loss of strategic opportunities in utilizing international payment and exchange capacities. In the context of Iran's sanctions, this is regarded as a strategic opportunity cost, highlighting the failure to leverage cryptocurrencies to counteract sanctions and achieve the resistance economy goals outlined in the Transformation Document. Finally, the high energy consumption resulting from unauthorized mining, with a weight of 0.114, ranked fourth. Although this is a significant infrastructural issue, its technical nature and higher controllability through policy tools place it lower in resource allocation priority.

Figure 1.
Output of the Expert Choice Software for the Main Challenges



(Source: The Researcher's Findings)

Table 2.
Ranking of the Main Challenges in Cryptocurrency Development

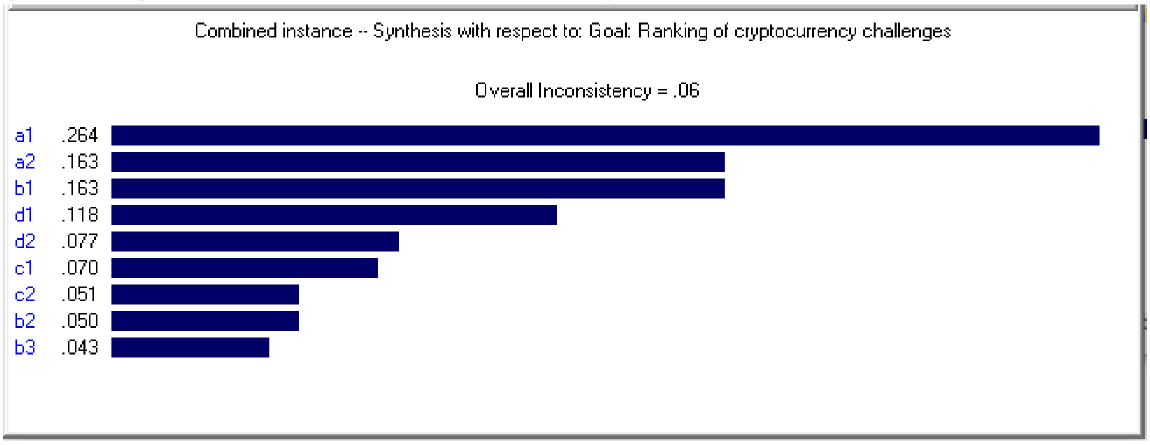
Row	Main challenges	Symbol	Coefficient
1	Weakness in macro-level management of cryptocurrencies	A	0.429
2	High risk to people's investment in the cryptocurrency market	B	0.264
3	Loss of opportunities in leveraging the strategic capacities of cryptocurrencies in the domestic and international payment and exchange domains	D	0.193
4	Rising share of hidden/mined extraction in the total cryptocurrency production markets in the country	C	0.114

(Source: The Researcher's Findings)

5.1 The Sub-criteria for Cryptocurrency Development Challenges

Analysis of the sub-criteria weights, conducted using AHP and Expert Choice software, provided an operational and precise prioritization of challenges in the cryptocurrency domain. This analysis is methodologically robust in two key aspects: first, its inconsistency rate of 0.06 falls below the standard threshold of 0.1; second, the primary criterion for prioritization has been the degree of alignment with the overarching objectives of the Transformation Document of the People's Government. In this ranking, "Inefficient governance structure in actively engaging with cryptocurrencies" held the top position with a weight of 0.264. This result clearly indicated that the core issue is the absence of a cohesive and proactive command structure, which both impedes effective governance and limits the utilization of existing opportunities. Following this, "Deficiencies in market monitoring, control, and supervision" and "Public's emotional motivation to escape inflation and invest in high-risk assets" both shared a weight of 0.163. The concentration of weights in the higher ranks of this analysis demonstrated that the government's immediate priorities should focus on two main areas of reforming the governance structure and managing socio-economic risks. In contrast, technical challenges such as "Lack of an intelligent electricity consumption monitoring framework" (ranked sixth with a weight of 0.070) and "Inappropriate energy pricing" (ranked seventh with a weight of 0.051) occupied lower positions. This weight distribution clearly indicated that, from the experts' perspective, governance and socio-economic issues carry significantly greater importance and urgency compared to purely infrastructural challenges.

Figure 2.
Output of the Expert Choice Software for Sub-criteria Challenges



(Source: The Researcher's Findings)

Table 3.

Ranking of the Sub-criteria/Challenges in Cryptocurrency Development

Row	Sub-indicators	Symbol	Coefficient
1	Inefficient governance structure in actively confronting cryptocurrencies; distributed ledger technology offices	a1	0.264
2	Weakness in monitoring, controlling, supervising,, and self- regulating the market for producing, storing, and exchanging cryptocurrencies	a2	0.163
3	Motivation and excitement of people to escape inflation and purchase non-Rial assets with high liquidity	b1	0.163
4	Liberation and ambiguity/uncertainty of cryptocurrency exchange and payment infrastructures domestically	d1	0.118
5	Excessive imports and capital outflow, and non-activation of exports	d2	0.077
6	Loss of an intelligent monitoring and control structure for electricity consumption patterns	c1	0.07
7	Inappropriate pricing of energy carriers and discrimination in dealing with the cryptocurrency mining industry compared to other industries	c2	0.051
8	Inadequate regulation in the area of asset management policies and custody of cryptocurrencies	b2	0.05
9	Weakness of policies supporting domestic production and export within the cryptocurrency ecosystem	b3	0.043

(Source: The Researcher's Findings)

Table 4.

Final Weights of the Challenges Cryptocurrency Development

Ranking of Cryptocurrency Development Challenges Based on the Government of the People Transformation Document						
Row	Main challenges	Coefficient	Row	Factors	Coefficient	rank
1	Weakness in macro-level management of cryptocurrencies	0.429	1	Inefficient governance structure in actively confronting cryptocurrencies; distributed ledger technology offices	0.264	First
			2	Weakness in monitoring, controlling, supervising,, and self-regulating the market for producing, storing, and exchanging cryptocurrencies	0.163	Second
2	High risk to people's investment in the cryptocurrency market	0.264	1	Motivation and excitement of people to escape inflation and purchase non-Rial assets with high liquidity	0.163	Third
			2	Inadequate regulation in the area of asset management policies and custody of cryptocurrencies	0.05	eighth
			3	Weakness of policies supporting domestic production and export within the cryptocurrency ecosystem	0.043	ninth
3	Rising share of hidden/mined extraction in the total cryptocurrency production market in the country	0.114	1	Loss of an intelligent monitoring and control structure for electricity consumption patterns	0.07	sixth
			2	Inappropriate pricing of energy carriers and discrimination in dealing with the cryptocurrency mining industry compared to other industries	0.051	seventh
4	Loss of opportunities in leveraging the strategic capacities of cryptocurrencies in the domestic and international payment and exchange domains	0.193	1	Liberation and ambiguity/uncertainty of cryptocurrency exchange and payment infrastructures domestically	0.118	Fourth
			2	Excessive imports and capital outflow, and non-activation of exports	0.077	fifth

(Source: The Researcher's Findings)

Sub-criteria Challenges	
MIN	MAX
0.043	0.264

Main Challenges	
MIN	MAX
0.114	0.429

(Source: The Researcher's Findings)

The final results AHP provided a decisive and well-founded roadmap for prioritizing policymaking in the cryptocurrency domain. The most critical challenge identified was the weakness in macro-level cryptocurrency management, with a significant weight of 0.429. This outcome indicated that the most fundamental threat in this field lies not in financial or technical issues, but in governance gaps and a passive decision-making structure. This inefficiency directly contradicted the proactive and opportunity-driven approach outlined in the Transformation Document of the People's Government. In contrast, the increase in hidden mining share, with a weight of 0.114, was identified as the least significant main challenge. Its technical nature and higher controllability have placed it at the bottom of the priority ranking. This prioritization is confirmed at the sub-factor level as well: the inefficient governance structure, with a weight of 0.264, is recognized as the most important sub-factor, emphasizing that without addressing this structural deficiency, any efforts to resolve other issues would prove futile. Conversely, weak policies supporting domestic production and exports, with a negligible weight of 0.043, was determined to be the least significant sub-factor. In summary, the final results clearly demonstrated that policymakers' urgent priority should focus on reforming the governance structure and managing socio-economic risks (stemming from public capital flight from inflation). This dual focus will enable both overcoming structural threats and effectively achieving the overarching objectives and international opportunities outlined in the Transformation Document of the People's Government

6. Discussion and Conclusion

The final results of AHP, validated by a strong and credible inconsistency rate of 0.06, provided a decisive roadmap and vital strategic insight for prioritizing policymaking in the cryptocurrency domain. While a broad consensus in the research literature globally (e.g., Adhikari et al., 2025) and domestically (Sadeghi et al., 2024) emphasized the existence of legal gaps, fraud risks, and technical challenges, the findings of the present study shifted the problem's center of gravity, revealing it at the institutional level. The decisive weight of 0.429 for weakness in macro-level cryptocurrency management and 0.264 for the sub-factor of inefficient governance structure indicated that the most fundamental obstacle is not technical in nature, but lies in the absence of a cohesive, active, and empowered command structure for exercising sovereignty. This outcome represented the core innovation of the research: the root of all operational risks (such as weak market monitoring with a weight of 0.163) is ultimately the direct consequence of the failure to implement the proactive and opportunity-driven approach to risk management outlined in theGovernment Transformation Document, highlighting a serious gap between the high-level goals of policymakers and their executive capabilities. Furthermore, the analysis of the final weights demonstrated that the policymaking priority must be focused on internal socio-economic risks; While global concerns (e.g.,Stoll et al., 2019) concentrated on environmental issues stemming from energy consumption, our results emphasized that the public's emotional motivation to escape inflation (with a final weight of 0.163) is of significantly higher policy urgency than the technical and infrastructural challenges related to energy (such as the lack of a smart electricity consumption monitoring framework with a final weight of 0.070). This finding confirmed that the threat against economic stability and public trust (key objectives of the Transformation Document) is far more critical for policymakers than infrastructural concerns, making the management of inflationary expectations a key issue in the cryptocurrency domain. In final conclusion, this research asserted that cryptocurrency development is a strategic opportunity to elevate Iran's position in the digital economy and counter sanctions (see Babazadeh et al., 2021 ; Zare, 2022) however, the effective attainment of the overarching goals and international opportunities enshrined in the Transformation Document is conditional upon the swift implementation of institutional reforms at the macro level. Consequently, the policymakers' immediate priority must focus on remedying the structural governance deficit and managing socio-economic risks. This is because of the fact that any effort to implement minor or technical policies without addressing the structural crisis ranked first will only lead to the wastage of resources, the repetition of inefficiencies, and the definitive loss of strategic opportunities. Therefore, this article, furnished a decisive tool for guiding resources towards the correct priorities by providing aweighted and credible hierarchy based on the highest executive document of thecountry.

Comparison with International Priorities: In sanctioned economies like Russia, cryptocurrency policies prioritize institutional control to evade sanctions and bolster monetary sovereignty, focusing on macro-level governance over technical issues like energy efficiency in mining (Hudima et al., 2022). Similarly, emerging markets show that cryptocurrency adoption correlates positively with regulatory quality but negatively with corruption, highlighting the need for robust institutional reforms to address fiscal risks before technical integration (Copestake et al., 2023). The IMF notes that in such contexts, "robust domestic institutions are critical to mitigate crypto-induced fiscal risks", as weak governance exacerbates economic instability under external pressures (Hacibedel & Perez-Saiz, 2023).In contrast, advanced economies like the EU and US leverage established institutions to prioritize financial stability and consumer protection over governance restructuring (Zetzsche et al., 2021). The EU's Markets in Crypto-Assets Regulation (MiCA), effective in 2024, mandate licensing and transparency for crypto-asset providers to ensure market integrity and investor safeguards (Huang et al., 2024). In the US, 2025 policies focus on SEC/CFTC oversight for anti-money laundering compliance, addressing operational risks like fraud within a stable institutional framework (Ordekian et al., 2025). These contrasts show that political-economic contexts shape cryptocurrency policy, with sanctioned economies addressing institutional voids and advanced jurisdictions focusing on regulatory perimeters for operational threats(Copestake et al., 2023) ; (Zetzsche et al., 2021).

Research Limitations

1. **Small Sample Size (N=13):** Although this sample size is relatively common and accepted in specialized multi-criteria decision-making studies, the limited sample size

means that the findings reflect only the views of a small group of experts and lack generalizability to a larger population (Adewumi et al., 2023).

2. **Subjective Bias:** The AHP method is a theory of measurement based on the subjective judgments and pairwise comparisons of experts. Although AHP includes mechanisms for measuring judgment inconsistency, the reliance on a scale of absolute judgments leads to a degree of subjectivism that can be influenced by the personal experiences of the experts (Saaty, 2008).
3. **Evolving Nature of Cryptocurrencies:** The results represented a cross-sectional snapshot of priorities at a specific point in time. Given the highly volatile nature of the cryptocurrency market and its susceptibility to speculative bubbles (Cheah & Fry, 2015), these priorities and relative weights can quickly shift in the near future with changes in technology or regulation both in Iran and globally.

Policy Recommendationstabli

Table 5.
Policy Recommendations

Row	Challenge	Policy Action	Expected Outcomes
1	Weakness in Macro-Level Management of Cryptocurrencies	Immediately establish or designate a single high-level entity with full executive authority for policymaking, regulation, and supervision over the entire cryptocurrency domain, adopting an opportunity-driven and proactive strategy.	Enhanced coordination and efficiency in macro-level cryptocurrency management.
		Clarify and separate responsibilities among entities such as the Central Bank, economic and security ministries, and IT organizations, replacing the inefficient decision-making structure with a cohesive chain of command.	Increased speed and accuracy in decision-making and policy implementation.
2	High Risk of Public Investment in the Cryptocurrency Market	Urgently develop a regulatory framework to minimize risks for small public investments and prevent loss of trust in the financial system.	Protection of small capital and increased public trust in the financial system.
		Organize domestic cryptocurrency exchange platforms with high transparency standards and launch a comprehensive public awareness campaign about the high-risk nature of these assets.	Greater public awareness and reduced reckless investments in high-risk markets.
3	Loss of Opportunities in Leveraging Strategic Capacities of Cryptocurrencies in Payment and Domestic/International Exchange	Swiftly shift policymaking to facilitate the legitimate use of cryptocurrencies in foreign trade and activate their potential to counter sanctions, in close collaboration with major exporters and importers.	Strengthened foreign trade and reduced impact of sanctions.
		Focus policymaking and regulation on the use of cryptocurrencies at the B2B level, rather than individual transactions, to ensure repatriation of foreign exchange earnings from exports.	Increased repatriation of foreign exchange earnings and bolstered national economy.
4	Rising Share of Hidden Mining in the Total Cryptocurrency Production Market	Immediately implement policies to legalize and regulate mining during times and in locations with high electricity production capacity, while disconnecting power to unauthorized mining farms.	Reduced illegal mining and optimized energy consumption.
		Allocate resources to this area without hindering the immediate and full implementation of reforms in higher-priority sectors (e.g., governance and public risk management).	Strengthened oversight and resource allocation to higher-priority sectors.

(Source: The Researcher's Findings)

REFERENCE

Adewumi, A. S., Onyango, V., Moyo, D., Al Waer, H., & Dawodu, A. (2023). Conceptualising the characteristics of the indicators of a neighbourhood sustainability assessment framework in a developing country context. *Environmental Impact Assessment Review, 102*, 107197. Retrieved from https://doi.org/10.1016/j.eiar.2023.107197.

Adhikari, P., Hamal, P., Adhikari, B., & Maskey, N. K. (2025). Cryptocurrency taxation and regulatory challenges. *International Journal of Science and Research Archive, 3*(15), 778–788. Retrieved from https://doi.org/10.30574/ijsra.2025.15.3.1791.

Ahmadpour, M. (2024). Analysis of threats and opportunities of cryptocurrencies on economic security as a component of national security. [In Persian]. *Journal of Monetary & Banking Research, 16*(58), 551–577. Retrieved from https://jmbr.mbri.ac.ir/article-1-2003-en.pdf

Ammous, S. (2018). Can cryptocurrencies fulfil the functions of money? *The Quarterly Review of Economics and Finance, 70*, 38–51. Retrieved from https://doi.org/10.1016/j.qref.2018.05.010.

Aref, M., Jamali, j., & Foroughnejad, H. (2024). Presenting a risk management model for The use of cryptocurrencies in Iran. [In Persian]. *Financial Knowledge of Securities Analysis, 62*(17), . Retrieved from http://sanad.iau.ir/fa/Article/802908.

Babak, M.,Hajimolamirzaee, H., & Najafi Jeze, H. (2024). The dimensions of crypto currencies policy making model in Islamic Republic of Iran. *Quarterly Journal of Interdisciplinary Studies on Strategic Knowledge, 14*(56), 245–223. Retrieved from https://smsnds.sndu.ac.ir/article_3316_8522f9adbffbad8eeed8f2d518f939c2.pdf

Babazadeh, Y., Farahmand, F.-h., Pasebani, M., & Alavi Matin, Y. (2021). A cnceptual model indicators facilitating the use of cryptocurrencies in international transactions in sanction situations. [in Persian]. *Journal of International Business Administration, 4*(1), 167–188. Retrieved from https://sid.ir/paper/954758/fa.

Bagban, E. (2019). *Legal challenges of Bitcoin activity in Iran*. Paper presented at the the Second National Cyber Defense Conference. https://civilica.com/doc/903723.

Cheah, E.-T., & Fry, J. (2015). Speculative bubbles in Bitcoin markets? An empirical investigation into the fundamental value of Bitcoin. *Economics letters, 130*, 32–36. Retrieved from https://doi.org/10.1016/j.econlet.2015.02.029.

Chen, W., & Phelan, G. (2025). Digital currency and banking-sector stability. *Journal of Financial Stability*, 101414. Retrieved from https://doi.org/10.1016/j.jfs.2025.101414.

Copestake, A., Furceri, D., & Gonzalez-Dominguez, P. (2023). Crypto market responses to digital asset policies. *Economics letters, 222*, 110949. Retrieved from https://doi.org/10.1016/j.econlet.2022.110949.

Corbet, S., Lucey, B., Urquhart, A., & Yarovaya, L. (2019). Cryptocurrencies as a financial asset: A systematic analysis. *International Review of Financial Analysis, 62*, 182–199. Retrieved from https://doi.org/10.1016/j.irfa.2018.09.003.

Ghaemi Asl, M., Abolhasani, S., & Farhoud, N. (2024). Prevention of criminal activity related to virtual currencies in Iran's criminal policy in line with the management of virtual currency markets. [In Persian]. *Journal of Islamic Marketing Research, 2*(3), 27–43. Retrieved from https://imr.yazd.iau.ir/article_709097_d3fb6bec09e1cf0da47697c7d275b3fa.pdf

HabibianNaqibi, M., Asadian Lalehmi, M., & Zahmatkesh, M. A. (2020). Examining the challenge of Gharar in cryptocurrency transactions (A case s Study of Bitcoin). [in Persian]. *International Conference on Jurisprudence, Law, and Religious Studies*. Retrieved from https://sid.ir/paper/900566/fa.

Hacibedel, M. B., & Perez-Saiz, H. (2023). *Assessing macrofinancial risks from crypto assets*: International Monetary Fund.

Hashem Nejad, N., Barkhordari, S., & Abdoli, G. (2024). Examining behavioral similarities between the cryptocurrency market and traditional financial markets (A case study of Bitcoin). [In Persian]. Retrieved from https://civilica.com/doc/2156112

He , Z., Li , Z., & Zhou, E. (2024). The Impact of cryptocurrencies and their regulatory challenges and countermeasures. *Highlights in Business, Economics and Management 45*, 978–983. Retrieved from https://www.researchgate.net/publication/388163795_The_Impact_of_Cryptocurrencies_and_Their_Regulatory_Challenges_and_Countermeasures.

Huang, W., Strack, P., & Tamuz, O. (2024). Learning in repeated interactions on networks. Econometrica, 92(1), 1-27. https://doi.org/10.3982/ECTA20806

Hudima, T., Dzhabrailov, R., & Chernykh, O. (2022).International economic situations against the Russian federation and cryptocurrencies: Legal assessment. *Entrepreneurship, Economy and Law* 27–36. Retrieved from https://www.researchgate.net/publication/368856745_INTERNATIONAL_ECONOMIC_SANCTIONS_AGAINST_THE_RUSSIAN_FEDERATION_AND_CRYPTOCURRENCIES_LEGAL_ASSESSMENT.

Fathi, M. R., Sadeghi, M. R., Ghadimi, S., & Akhlaghpour, S. (2025). Identifying and ranking barriers to implementing Internet of things in food supply chains: A case study of Kalleh company. Knowledge Economy Studies, 2(1), 137-156. 10.22034/kes.2025.2057601.1057.

Iran, T. P. o. t. I. R. o. (2022). People's government transformation document. [In Persian]. *The National System of Laws and Regulations of the Islamic Republic of Iran*. Retrieved from https://qavanin.ir/Law/TreeText/?IDS=15184881572116350348.

Kashian, A., & Parnian, N. (2021). Assessing the impact of crypto currencies on the spread of money laundering in Iran. [In Persian]. *Karagah, 15*(54), 83–104. Retrieved from https://sid.ir/paper/961461/fa.

Kerr, D. S., Loveland, K. A., Smith, K. T., & Smith, L. M. (2023). Cryptocurrency risks, fraud cases, and financial performance. *Risks, 11*(3), 51. Retrieved from https://www.mdpi.com/2227-9091/11/3/51

Larijani, S., & Taheri, A. (2022). Assessing the economic feasibility of merging small-scale CHP power plants with cryptocurrency mining farms: Effects on energy supply security and passive defense of Iran. [In Persian]. *Defense Economics & Sustainable Development, 7*(25), 67–89. Retrieved from https://eghtesad.sndu.ac.ir/article_2429_eb3a2326585d3216a59963520290ca9a.pdf

Li, D., & Huang, J. (2019). The impact of digital currency on the financial system: Universal decentralized digital currency, is it possible? *Journal of Fintech, Blockchain, and Smart Contracts, 2*(3), 21–25. Retrieved from http://dx.doi.org/10.22158/jepf.v5n2p203

Lourie, B., Nekrasov, A., Truong, P., & Zhu, C. (2023). Crypto fraud and investing behavior. *Available at SSRN 4650849*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4650849

Mohammadi Fateh, A., & Salarnejad, A. A. (2022). The scope of blockchain technology: A meta-synthesis study of applications, benefits, challenges, and related technologies. [in Persian]. *Sciences and Techniques of Information Management, 8*(1), 245–300. Retrieved from https://stim.qom.ac.ir/article_2137_ea5910fb79d4be7448252639922e5164.pdf

Ordekian, M., Becker, I., Moore, T., & Vasek, M. (2025). Raising the bar: Assessing historical cryptocurrency exchange practices in light of the EU's MiCA and DORA regulation. *Available at SSRN 5327395*. Retrieved from https://papers.ssrn.com/sol3/Delivery.cfm?abstractid=5327395

Oye, E., Martins, S., & Anderson, G. (2025). Regulatory challenges in combating cryptocurrency-based money laundering. Retrieved from https://www.researchgate.net/profile/Emma-Oye/publication/394471722_Regulatory_Challenges_in_Combating_Cryptocurrency-Based_Money_Laundering/links/689d39f6495bc343ed4af834/Regulatory-Challenges-in-Combating-Cryptocurrency-Based-Money-Laundering.pdf

Rajabi, A., & Saberi, A. (2022). The impact of cryptocurrency mining on the sustainability of Iran's power grid. [In Persian]. *Research Center of the Islamic Consultative Assembly*. Retrieved from https://sid.ir/paper/1048297/fa

Rostami, A., ArabSorkhi, A., Aghaei, M., & Sokhai, M .A. (2024). Regulatory challenges of cryptocurrencies in the Islamic Republic of Iran. [in Persian]. *Quarterly Journal of Strategic Environmental Studies of the Islamic Republic of Iran, 8*(3), 79–110. Retrieved from http://jse.sndu.ac.ir/article_3086_6c4df20b43100ce3362b97cfb96d908b.pdf

Rostami, A., Jalalifarahanh, G., & Kavoosi, H. (2024). Identify the effects of using cryptocurrencies on the economic aspects of national security. [In Persian]. *Quarterly Journal of Interdisciplinary Studies on Strategic Knowledge, 14*(54), 155–133. Retrieved from https://smsnds.sndu.ac.ir/article_3026_f4e547cce52d1ee0f22566fff2dce512.pdf

Saaty, T. L. (2008). Decision making with the analytic hierarchy process. *International journal of services sciences, 1*(1), 83–98. Retrieved from https://doi.org/10.1504/IJSSCI.2008.017590

Sadeghi, M. R., Fathi, M. R., Khanjani, M. & Mansouri, M. (2025). Modeling the barriers to implementing Artificial Intelligence in desalination supply chain using MICMAC. *Knowledge Economy Studies*, *2*(2), 85-104. 10.22034/kes.2025.2069582.1071

Sadeghi, H., Nouri, F., & Naser, M. (2024). A comparative study in the analysis of the legal format of the cryptocurrency Technology mining mechanism: needs and policies. [in Persian]. *Roshd -e- Fanavari, 79*(20), 11–19. Retrieved from http://rimag.ir/fa/Article/40102

Saeidi, P., Jouybari, A. N., Khorasani, M. B., & Naderian, A. (2024). Designing the pattern of business intelligence application In the cryptocurrency market using grounded theory. [in Persian]. *Journal of Capital Market Analysis, 3*(4), 189–221. Retrieved from http://sanad.iau.ir/fa/Article/1127596

Shamsi, R., Abdulahi, E., & Safaei, M. (2024). The impact of cryptocurrencies on decision-making processes and the legal requirements for criminalizing fraud and forgery in this context. [In Persian]. *The Islamic Revolution Approach, 18*(68), 257–280. Retrieved from https://www.rahyaftjournal.ir/article_213180_5c044634b5560317ccb338acf09126b5.pdf

Stoll, C., Klaaßen, L., & Gallersdörfer, U. (2019). The carbon footprint of bitcoin. *Joule, 3*(7), 1647–1661. Retrieved from https://www.cell.com/joule/fulltext/S2542-4351(19)30255-7?source=post_page

Taj Langerudi, M. H., & Dehdar, F. (2024). The challenges of using cryptocurrencies in the legal system of Islamic Republic of Iran. [In Persian]. *ModernTechnologies Law, 5*(9), 87–106. 10.22133/mtlj.2023.401446.1216

Varmaziari, F., Saneiian, A., & Fathian, M. (2024). Strategic planning for the development of cryptocurrencies in the country's banking system. [in Persian]. *Organizational Resource*

Management Research،. Retrieved from https://civilica.com/doc/2152046

Zare, A. (2022). Investigating the position of cryptocurrencies in world trade with a look at the approach of international organizations and Iran. [in Persian]. *Modern Interdisciplinary Research in Law*. Retrieved from https://dorl.net/dor/20.1001.1.27833631.1401.2.3.6.0

Zetzsche, D. A., Annunziata, F., Arner, D. W., & Buckley, R. P. (2021). The Markets in Crypto-Assets regulation (MiCA) and the EU digital finance strategy. *Capital Markets Law Journal, 16*(2), 203–225. Retrieved from https://doi.org/10.1093/cmlj/kmab005

Zheng, Q. (2025). The risk impact of virtual currencies on financial market stability and preventive strategies. *Advances in Economics Management and Political Sciences, 199*(1), 124–131. Retrieved from https://www.researchgate.net/publication/393611702_The_Risk_Impact_of_Virtual_Currencies_on_Financial_Market_Stability_and_Preventive_Strategies

Zimba, A., Phiri, K. O., Mulenga, M., & Mukupa, G. (2025). Blockchain technology and energy efficiency: A systematic literature review of consensus mechanisms, architectural innovations, and sustainable solutions. Retrieved from https://doi.org/10.21203/rs.3.rs-6457924/v1

Zulhuda, S. (2017). Whither policing cryptocurrency in Malaysia? *IIUMLJ, 25*, 179. Retrieved from https://doi.org/10.31436/iiumlj.v25i2.342