ITS-SGBM: A Hybrid Intelligent Tuna Swarm Optimization and Stochastic Gradient Boosting Machine Model for Financial Performance Forecasting

Shixin Liang, Kang Zhang*
Office of Planning Finance, Hebei Petroleum University of Technology, 067000, China
E-mail: KangZhang2025@outlook.com, IsaiahMargaret33148@outlook.com
*Corresponding author

Keywords: decision-making models, financial performance, accounting systems, TOPSIS, intelligent tuna swarm optimized stochastic gradient boosting machine (ITS-SGBM)

Received: July 15, 2025

In the evolving financial landscape, precise performance measurement is crucial for informed decision-making in accounting systems. Conventional approaches often struggle to handle complex business processes and regulatory changes. The research suggests a hybrid Intelligent Tuna Swarm Optimized Stochastic Gradient Boosting Machine (ITS-SGBM) model for improving financial performance assessment. Financial data from balance sheets and annual reports (2009–2022) were pre-processed through normalization, and Principal Component Analysis (PCA) was extracted as features. The dataset was split into training and testing sets, and ITS-SGBM parameters were enhanced using Intelligent Tuna Swarm Optimization to balance exploration and exploitation during model training. Model performance was confirmed using k-fold cross-validation and benchmarked against remaining models, including DGRU-IMPA and DNN-OSIE-CHOA. In addition, the TOPSIS decision-making framework was functional to rank firms based on expert-defined financial criteria. Outcomes show that the offered model accomplished superior forecasting accuracy (96.2%) with lower error rates (MAE: 0.009, RMSE: 0.005, RRMSE: 0.072) and ranking consistency of 0.94, outperforming the benchmark models. The findings confirm that the ITS-SGBM–TOPSIS incorporation provides a strong, data-driven decision-support mechanism, increasing financial stability, strategic planning, and sustainable business development.

Povzetek: Raziskava predlaga hibridni model ITS-SGBM, ki z optimizacijo in metodo TOPSIS bistveno izboljša napovedovanje finančne uspešnosti ter razvrščanje podjetij, pri čemer doseže visoko natančnost in preseže primerjalne modele.

1 Introduction

A dynamic business climate demands financial performance evaluation to be a primary element for organizational decision-making [1]. Accounting systems are important in strategic planning in terms of financial evaluation of health, as they maintain the requirements of regulations. The assessment of financial performance with proper procedures enables companies to identify the level of profitability in the company as it manages its risks and enhances the allocation of its resources [2]. The need to enhance financial performance evaluation in accounting systems has become more evident due to its effects on both accuracy and power in strategic decision making [3].

The financial performance evaluation process broadens its activities beyond the corporate profitability goals. Organizations do not just provide financial reports merely to satisfy the business stability needs of their investors and stakeholders, but also their development insight needs [4]. The issues of financial analysis present scenarios leading to either erroneous stakeholder behavior or low performance and financial problems for the organization. The modern accounting systems that are capable of meeting modern business needs are composed of superior data analysis and real-time tracking features, accompanied by improved reporting systems. The enhanced reporting strategy will integrate the conventional accounting methods and the current business requirements to accommodate the complex economic situations [5].

The development of a high-quality accounting system in the assessment of financial performance is related to the long-term growth and creation of competitive advantage for the business [6]. The companies operating under more evaluated systems have a better ability to predict financial risks, make the most of their spending, and improve financial security. Besides that, the implementation of ESG-based assessment practices and green accounting is focused on sustainability and enables highlighting the importance of ensuring the appropriate balance between environmental and social responsibility and financial accuracy [7]. The accounting systems provide panoramic improvements that give precise and timely financial analyses, which are valuable both to individual companies and to the global industrial and economic well-being [8].

Existing models are large dataset-based and sensitive to parameter settings, which cause overfitting and poor generalization to small or noisy financial data. They are plagued by premature convergence, overly complex computation, and the ability to scale to dynamic financial conditions. Minimal research has been conducted to incorporate into multi-criteria decision-making models. ITS-SGBM is also effectively utilized in the optimization of parameters without premature convergence and has enhanced prediction accuracy when considering small or noisy data. SGBM lowers the overfitting using ensemble learning and regularization, thus increasing reliability. TOPSIS integration also facilitates expert-informed and scalable ranking of firms to make reliable data-driven and financial decisions.

This research aims to create a hybrid ITS-SGBM-TOPSIS architecture of precise and scalable financial performance assessment in accounting systems. It is used to increase the predictive accuracy of major financial measures and decrease errors and overfitting. It gives a good, reliable, informed firm ranking as a way of supporting data-based strategic decision-making and sustainable business development.

The rest of the research is separated into the subsequent parts: section 2 discusses the related research; section 3 outlines the proposed technique approach; section 4 presents the results and discussion; and section 5 summarizes the overall investigation's conclusion.

1.1 Key contribution

- The research presents a combined approach to the evaluation of financial performance in accounting systems, which improves the accuracy, strength, and dependability of the firm's ranking in the dynamism of financial markets.
- To collect data, between 2009 and 2022, balance sheets, income statements, and cash flow statements were used, and data were processed through normalization and extraction of Principal Component Analysis (PCA) to stabilize and use the data to predict reliably.
- It is a combination of Intelligent Tuna Swarm Optimization (ITS) and Stochastic Gradient Boosting Machine (SGBM) that enables optimized parameter tuning, exploration/exploitation, and reduces overfitting to ensure high predictive power and robustness across a broad spectrum of data.
- The findings reveal that the forecasting accuracy (96.2) is higher, as well as the error (MAE: 0.009, RMSE: 0.005, RRMSE: 0.072), and firm ranking (0.94) are better than the baseline models DGRU-IMPA and DNN-OSIE-CHOA, which support reliable, data-driven decision-making.
- The methodology offers a scalable system of decision support through the combination of TOPSIS and ITS-SGBM, which sustains strategic financial planning, risk management, and long-term sustainable business development in a complicated accounting system.

1.2 Research question

How effectively can the ITS-SGBM model predict financial performance metrics, such as total revenue, compared to existing models like DGRU-IMPA and DNN-OSIE-CHOA?

Can Intelligent Tuna Swarm Optimization improve model parameter tuning to enhance prediction accuracy and robustness across diverse and noisy financial datasets?

How does integrating ITS-SGBM with the TOPSIS decision-making framework improve the accuracy and reliability of firm rankings based on expert-defined financial criteria?

2 Related works

A framework of policy-maker decision-making based on a portfolio to focus on the Internet of Things (IoT) applications in agriculture was created to balance sustainability [9]. The methodology applied meta-synthesis, fuzzy Delphi, IVTFN-SWARA (Interval-Valued Triangular Fuzzy Number step-wise Weight Assessment Ratio Analysis). Findings indicate that the framework aids in making disciplined decisions regarding investing in IoT, but it is constrained by assumptions based on expert judgment and selected criteria, as well as validation under different real-life conditions. To improve decision-making with fuzzy and unreliable data, the Evaluation Based on Distance from Average Solution (EDAS) was developed [10]. The research applied a decision support system accounting for expert reliability, which was applied to wind energy investment for selecting the best wind turbine. The findings show sensible and consistent ranking and better agreement with expert judgements. However, they are also limited in their generalizability, for the method. The research proposed a new multiple-criteria decision-making (MCDM), MULTIMOOSRAL, Combined Compromise Solution (CoCoSo), a new ranking model that substitutes the usual dominance theory [11]. Results give a more precise and valid ranking of the alternatives, particularly when the performances are close. Nevertheless, it is only validated in terms of an illustrative case of a supplier selection. Creating a deep learning-based performance evaluation method was the objective of the exploration [12] to support the long-term growth of accounting companies. It increased the restricted Boltzmann machines (RBM) model's accuracy by employing deep belief networks

(DBN) and reverse fine-tuning. The purpose of the examination [13] was to examine financial accounting data to provide insights into performance optimization, cost reduction, and profitability enhancement. The purpose of the research [14] was to evaluate the performance of ML classifiers in identifying financial fraud in Turkish small and medium-sized enterprises (SMEs) by analyzing financial statements. A Fermatean Fuzzy ELECTRE (Elimination and Choice Expressing Reality) method for multi-criteria group decision-making was examined [15]. The study proposes a novel distance measure for Fermatean fuzzy sets based on Jensen–Shannon divergence, which is cross-entropy. The results indicate feasibility, and the comparative analysis indicates superior performance. However, the validation is limited to the case presented in the paper, and broader investigations of real-world scenarios are necessary. By utilizing a hybrid Twin Adjustable Reinforced Chimp Optimization Algorithm (TAR-CHOA) and Deep Long Short-Term Memory (DLSTM) model, the examination [16] seeks to enhance profit prediction in Financial Accounting Information Systems (FAISs). Table 1 illustrates the related articles' results and their limitations.

Table 1: Summarizes findings from related studies, highlighting their key results and associated findings.

Ref	Objective	Method	Result	Limitation
Mohammadian et al. [9]	To create a portfolio-based decision-making model to be used by policy-makers to prioritize IoT applications in agriculture, to balance sustainability and risk.	Applied meta-synthesis, fuzzy Delphi, IVTFN-SWARA (Interval-Valued Triangular Fuzzy Number Step-wise Weight Assessment Ratio Analysis), and IVTFN-ARAS in an MADM (Multi-Attribute Decision-Making) model.	The framework facilitates disciplined investment choices in the application of IoT and assists in prioritizing the areas with consideration of risk-return trade-offs.	Restricted due to the dependence on the expert judgment, the choice of criteria, and the absence of validation in a wider range of real-world situations.
Tüysüz & Kahraman [10]	Develop a fuzzy MCDM methodology that accounts for the reliability of expert evaluations under vague and imprecise data.	An Integrated Z-fuzzy Analytic Hierarchy Process (AHP) for criteria weighting combined with Z-fuzzy Evaluation Based on Distance from Average Solution (EDAS)	Provides rational, consistent rankings of alternatives, accurately reflecting expert reliability and enabling decision-making related to wind energy investment.	Applicability has only been verified in the selection of wind turbines; further validation in a wider range of sectors is necessary.
Ulutaş et al. [11]	To develop a reliable multiple-criteria decision-making (MCDM) method for selecting the most appropriate alternative	Proposed MULTIMOOSRAL method integrating CoCoSo and a logarithmic approximation with a novel ranking approach.	Gives better and more credible ranking of alternatives, especially when performances are similar.	Validation limited to an illustrative supplier selection case; broader applicability requires further testing.
Hu [12]	To design a deep learning-based method for sustainable accounting evaluation.	Applied Restricted Boltzmann Machines (RBM) enhanced with Deep Belief Networks (DBN) and reverse fine-tuning.	Achieved higher Average Precision (AP), Average Recall (AR), and overall accuracy compared to other approaches.	High dependence on large datasets and computational complexity reduces scalability.
Chakri et al. [13]	To analyze financial accounting data for	Conducted Exploratory Data Analysis (EDA) and applied a Decision Tree classifier	The Decision Tree with an optimal depth of nine	Limited generalization ability and sensitivity to data

	performance optimization.	optimized through grid search.	provided the best performance.	variability reduce robustness.
Hamal & Senvar [14]	To detect financial accounting fraud in small and medium-sized enterprises (SMEs).	Compared several Machine Learning (ML) classifiers, with oversampling applied to balance datasets.	Random Forest with oversampling achieved the highest fraud detection accuracy.	Possibility of data bias due to reliance on a single bank's dataset.
Zhou et al. [15]	To create a Fermatean fuzzy ELECTRE method for multi-criteria group decision-making with decision-maker weights unknown and criteria weights incomplete.	A new distance measure is suggested based on Jensen–Shannon divergence and cross-entropy for Fermatean fuzzy sets, dynamic determination of the decision-maker.	The feasibility and greater efficacy were shown in a real-world case study by comparative analysis.	Validity tested applicable only in the depicted case; further validity checks are warranted in different decision-making contexts.
Tang et al. [16]	To optimize profit prediction in accounting information systems.	Combined Deep Long Short-Term Memory (DLSTM) with the Twin Adjustable Reinforced Chimp Optimization Algorithm (TAR-CHOA).	DLSTM–TAR-CHOA achieved superior profit prediction accuracy compared to alternatives.	Strong reliance on historical data and sensitivity to market volatility reduced model robustness.

2.1 Research gap

The decision model of the IoT portfolio is based on expert judgement and has not been proven in the wider applications of real-life applications [9]. The fuzzy MCDM method guarantees credible rankings but has been experimented on wind energy investment only, and it has to be extended to broader uses [10]. The MULTIMOOSRAL approach enhances the ranking credibility, but its validation is limited to suppliers' selection, whereas the deep learning-based accounting framework has a scale problem related to computational limitations [11- 12]. Furthermore, the real-time adaptive performance and integration with multi-criteria decision frameworks are poorly studied. The present research addresses these shortcomings by suggesting a hybrid ITS-SGBM-TOPSIS model that is highly accurate and scalable, with sound decision-making in diverse financial settings.

3 Methodology

The process includes extracting financial data from balance sheets, income statements, and cash flow statements, followed by key financial ratio computation and ITS-SGBM-based revenue forecasting. The financial performance measurement of TOPSIS occurs through expert-established criteria that create its rankings. Figure 1 illustrates the proposed research's basic concept.

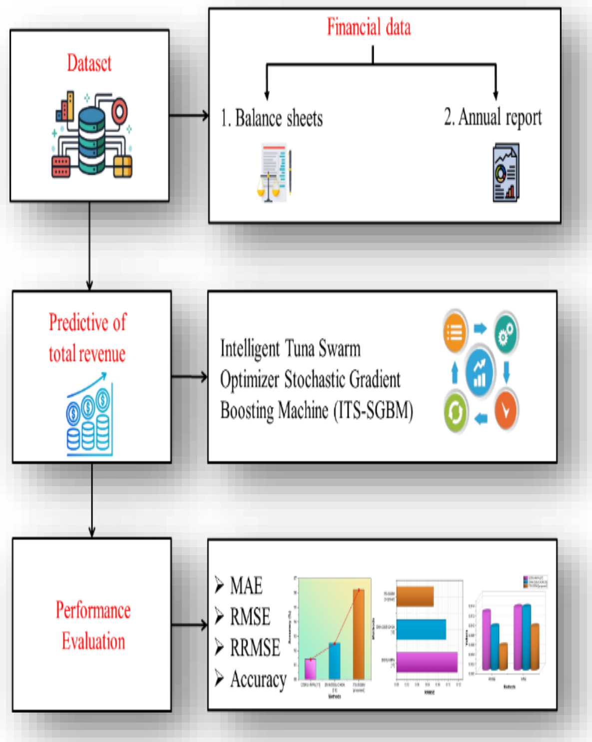

Figure 1: Fundamental concept of proposed research

A. Dataset

The accumulation of data from balance sheets and 10-K annual reports from different companies. The data, which spans the years 2009–2023, is longitudinal or panel data. A few insolvent enterprises are included to aid in the investigation of contributing causes. The firms' names are listed based on their stock prices. Companies were separated into distinct groups. The data was split into training and testing, 80% and 20% respectively, and further validated model-performance-wise by the cross-validation method to improve robustness and to reduce any possible overfitting. The data gathered from Kaggle: https://www.kaggle.com/datasets/rish59/financial-statements-of-major-companies2009-2023

3.2 Preprocessing

It cleanses, scales, and transforms features of financial data, making it consistent and less noisy. This is facilitated because it allows the ITS-SGBM model to learn successfully and generate precise and reliable projections of financial performance.

3.2.1. Data preprocessing using Min-max normalization

The normalization is used to standardize all the financial features such that there are no dominant features on the basis of their magnitude, and it learns in a stable and unbiased way. It is to create a hybrid ITS-SGBM-TOPSIS framework for the true, robust, and scalable evaluation of financial performance in accounting systems. To improve the accuracy of the prediction of important financial ratios, minimize errors, and offer credible company rankings that are used in the decision-making process based on data. It is described as equation (1)

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{1}$$

Where:

- x'= normalized value
- x = original value
- x_{max} = maximum value in the dataset
- x_{min} = minimum value in the dataset

This preprocessing step normalizes the data and attenuates noise, facilitating stable and high-quality learning outcomes in the ITS-SGBM model.

3.3 Feature extraction using Principal Component Analysis (PCA)

PCA was used on the pre-processed annual report data and balance sheet of the companies to enhance financial performance forecasting and eliminate redundancy in highly correlated financial indicators. PCA converts the initial correlated features into a reduced number of uncorrelated elements (principal components) and leaves the largest amount of variance in the data. It augmenting the predictive power of the ITS-SGBM model by deriving the most informative attributes and reducing noise and multicollinearity. It is formulated as equation (2)

$$Z = XW \tag{2}$$

Where Z= normalized feature matrix, X= matrix of eigenvectors of the covariance matrix of X W= principal components representing transformed, uncorrelated features. The principal components Z are then used as input for ITS-SGBM, enabling more accurate and efficient financial performance prediction. PCA was effective in deriving the most informative financial indicators as it had been reduced to a smaller number of dimensions, which improved the predictive power of the ITS-SGBM model.

1.4. ITS-SGBM uses a predictive total revenue

The ITS-SGBM hybrid template is a combination of ITS and SGBM in the financial performance analysis. ITS guarantees the optimization of parameters through an efficient exploration and exploitation. SGBM enhances predictive power by the application of ensemble learning and regularization, which reduces overfitting. It is a powerful and scalable finance forecasting and decision-making system that is provided by this hybrid framework.

1.4.1. Stochastic Gradient Boosting (SGB)

SGB is used to extract complex financial trends using historical information and enhance stability in prediction. It adds randomness in the sampling to enhance generalization, and thus, the financial performance assessment becomes more adaptive and reliable. To predict the entire income of a company, the SGBM uses machine learning algorithms. It uses historical data on finances. By increasing forecast accuracy, SGB makes financial evaluations more trustworthy and assists companies in making data-driven, strategic decisions for stability and development. SGBM is an improvement of the standard Gradient Boosting (GB) algorithm by Friedman, where model verification through iterative learning is allowed. With a training sample $\left(w_j, z_j\right)_{j=1}^{M}$, and a randomly permuted set of indices $(\pi\ (j))_{j=1}^{M}$, the SGBM algorithm follows equations (3-4).

Initialized model: $E_0(w) = median\{z_j\}_{j=1}^{M}$

(3)

For $\mathrm{n} = 1$ to N, iterate

Shuffled data: $\{\pi(j)\}_1^M = rand\_perm\ \{j\}_j^M,$

(4)

Where $E_0(w)$ is the initial prediction set as the median of target values, z_j represents the actual financial outcome for each training instance and M is the total number of samples. $\pi(j)$ denotes the randomly permuted indices in each iteration, allowing the model to learn generalized patterns and improve forecasting accuracy. The randomized subsample of size $\sim M < M$ is denoted by $\{\bar{z}_{\pi(j)n}, w_{\pi(j)}\}_j^{\sim} M$, and M is the number of trees and components of the original training set, respectively. The following equation (5) is the definition of the proportion of patterns utilized in this stochastic model to train the base learner:

$e = \frac{\sim M}{M}.$

(5)

Where e represents the normalized error for the model, $\sim M$ is the cumulative prediction error across all training instances, and M is the total number of samples. The procedure is equivalent to the deterministic variant if e has a value of 1, which can range from 0.5 to 1. To lessen the over-fitting effect, Friedman suggests using a value of $e = 0.5$. $e = 0.8$ produced the best performance in the experiments it conducted. This SGBM enhances prediction efficiency and computation velocity, which makes it appropriate for application in forecasting in accounting systems.

1.4.2. Intelligent tuna swarm (ITS) optimization

ITS optimization is used to optimize the model parameters by balancing between exploration and exploitation during training. It eliminates premature convergence and enhances search performance, making the financial performance forecasting more robust. To estimate financial performance accurately, the SGBM is improved using ITS. Through model parameter optimization, ITS increases the accuracy of revenue forecasts.

3.4.2.1 Initialization of a piecewise chaotic map

Chaos theory optimizes better by ensuring distributed and diverse solutions. To enhance financial performance prediction, it applies Piecewise Chaotic Map Initialization to initialize the positions of financial points. It prevents random clustering and enhances search efficiency. The chaotic map equation (6) is:

$$\mathrm{w}_{\mathrm{l}+1} = \begin{cases} \mathrm{W_l}/\mathrm{O} & 0 \le \mathrm{W_l} < \mathrm{O} \\ (\mathrm{w_l} - \mathrm{o})/(0.5 - \mathrm{O}) & 0 \le \mathrm{W_l} < 0.5 \\ (1 - \mathrm{O} - \mathrm{W_l})/(0.5 - \mathrm{O}) & 0.5 \le \mathrm{W_l} < 1 - \mathrm{O} \\ (1 - \mathrm{W_l})/\mathrm{O} & 1 - \mathrm{O} \le \mathrm{W_l} < 1 \end{cases}$$

(6)

Where O and $\mathrm{W_l}$ are between [0, 1]. This method enhances the prediction of financial trends by providing a better-balanced model training dataset.

1.4.2.1. Optimal search using cubic map fusion

To enhance the prediction performance, Cubic Map Fusion with Inverse Refraction Initialization is used in place of random initialization. This assists in balancing the exploitation and exploration for financial predictions. The cubic chaos function is the following equation (7):

$$z_{j+1} = 2z_j^3 - \frac{3z_j}{2}, \quad -1 < z_j < 1,\ z_j \neq 0$$

(7)

Where z_j represents the cubic chaos function. It generates a chaotic sequence to improve exploration and exploitation, enhancing the search for optimal parameters in financial predictions. The updated tuna locations are plotted as follows:

$$W_j = \frac{(W_{va} - W_{ka})(z_i + 1)}{2} + W_{kb}$$

(8)

In Equation (10), W_j represents the mapped tuna position, whereas W_{va} and W_{ka} are the maximum and minimum thresholds in the computation space, and W_{kb} is an adjustment factor. In addition, Refractive Backward Learning and Opposition-Based Learning (OBL) improve solution quality by investigating the opposite trend of financial drivers. The optimization equation (9) is:

$$W_{ji} = (b_i + a_i)/2 \ + (b_i + a_i)/2l - W_{ji}/l$$

(9)

Where l improves the OBL process, a_i and b_i are the lower and upper bounds of the search space for the i parameter, and j is the current iteration index, Wji representing the current weight of the parameter. This enhances the accuracy of financial predictions by eliminating local optima.

1.4.3. Technique for order of preference by similarity to ideal solution (TOPSIS)

Decision-making through the TOPSIS methodology provides an effective solution for expert-based ranking of financial performance. The TOPSIS method determines financial alternative values by assessing their closeness to the Positive Ideal Solution (PIS) as well as the Negative Ideal Solution (NIS). The financial criteria weights in the TOPSIS model were calculated using a Delphi-based expert elicitation process. Five senior financial analysts and two academic gurus in corporate finance engaged in two rounds of consultations that were conducted in a manner of repetition. It was requested to give relative importance scores of each of these criteria (Liquidity Ratio, Leverage Ratio, Coverage Ratio, and Profitability Ratio) under industry best practices and their best judgment. The normalized scores were averaged to derive the final weights; this way, the weighting scheme derived is consistent with both practice and academics. By using this methodology, financial entities receive objective placement rankings according to their financial stability and health. The process follows these steps:

Phase 1: Clearly define the ITS-SGMB problem, including its standards and potential. Then, create a decision matrix that assesses the potential in light of the standards.

Phase 2: Employing Equation (10), the decision matrix is normalized.

$$Q = \left(q_{ji}\right)_{n\times m} = {w_{ji}}\Big/\left(\sqrt{\sum_{j=1}^{n} w_{ji}^2}\right) \tag{10}$$

Where q_{ji} offers the normalized value of the i^{th} financial criterion for the j^{th} company, w_{ji} remains the original score of that criterion, n signifies the total number of companies, and m stands for the total number of financial criteria.

Phase 3: Use Equation (11) to create a weighted standardized decision matrix.

$$U = \left(u_{ji}\right)_{n\times m} = x_i \times q_{ji} \tag{11}$$

Where u_{ji} remains the weighted normalized value of the i criterion for the j company, and x_i produces the expert-assigned weight of the i financial criterion.

Phase 4: Determine the PIS and NIS by applying equations (12).

$$B^+ = \{u_1^+, u_2^+, \dots, u_m^+\} \tag{12}$$

Here u_1^+ represents the Positive Ideal Solution (PIS) for the m criterion. B^+ takes the maximum weighted value.

Phase 5: Use Equations (13) to determine the distance between each option and M and the PIS to evaluate how effectively potential j performs.

$$c_j^+ = \sqrt{\sum_{i=1}^{n}\left(U_j - U_i^+\right)^2} \tag{13}$$

Where c_j^+ describes the Euclidean distance of the j company from the PIS, U_j are the weighted normalized value for the j company, and U_i^+ is the PIS value for the i criterion.

Step 6: To rank the options, each one's closeness coefficient is determined using Equation (14).

$$dd_j = \frac{c_j^-}{c_j^+ - c_j^-} \tag{14}$$

where d is the closeness coefficient of the j company, and c_j^- stands the distance to the Negative Ideal Solution (NIS). A d_j higher indicates better overall financial performance, allowing ranking of companies objectively.

The exploration incorporates TOPSIS into financial ranking to deliver a data-based system that evaluates companies through essential financial indicators. The expert-established weights for financial criteria help the rankings comply with industrial best practices, which results in more dependable financial decisions. Algorithm 1 shows the process of ITS-SGBM

Algorithm 1: Intelligent Tuna Swarm Optimized Stochastic Gradient Boosting Machine (ITS-SGBM)

Input: Financial dataset D, decision matrix W, population size P, iterations T
Output: Predicted Revenue, Firm Ranking
1. Initialize SGBM with base learner E0(w) = median(y)
2. Use ITS to optimize SGBM parameters:
- Initialize tuna positions with a chaotic map
- Update positions with cubic map + OBL
- Evaluate fitness = RMSE (SGBM (D, params))
- Keep best W_{best}

3. Train final SGBM with optimized $W_{best} \rightarrow$ predict revenue
4. Apply TOPSIS on decision matrix W:
   - Normalize and weight matrix
   - Find PIS and NIS
   - Compute distances $cj+, cj-$
   - Compute closeness $dj = cj-/(cj+ + cj-)$
   - Rank firms by dj
5. Return Predicted Revenue, Ranked Firms

4 Experimental setup

Table 2 shows the system setup and the tools that were used in the ITS-SGBM-TOPSIS financial performance forecasting framework.

Table 2: Experimental Setup for ITS-SGBM–TOPSIS framework

Component	Specification
OS	Ubuntu 22.04 LTS
CPU	Intel Core i7-12700K, 12-core @ 3.6–5.0 GHz
RAM	32 GB DDR4
GPU	NVIDIA RTX 3080 (10 GB)
Python Version	3.9
Frameworks	CatBoost, ITS-SGBM
Optimization Tools	Intelligent Tuna Swarm Optimization (ITS), Cubic Map Fusion
Feature Extraction	PCA
Classifier/Model	ITS-SGBM
Decision-Making	TOPSIS
Metrics	MAE, RMSE, RRMSE, Forecast Accuracy, Ranking Consistency
Visualization Tools	Matplotlib, Seaborn

4.1 Parameters setup

Table 3 explains the hyperparameters of the ITS-SGBM model.

Table 3: Hyperparameter Setup for ITS-SGBM

Hyperparameters	Typical Values
Number of Trees	100, 200, 300
Learning Rate	0.01, 0.05, 0.1
Maximum Depth	3, 5, 7
Minimum Samples per Leaf	1, 3, 5
Subsample Ratio	0.6, 0.8, 1.0
Column Subsample Ratio	0.6, 0.8, 1.0
Iterations of ITS Optimization	50, 100
Exploration–Exploitation Balance	0.3, 0.5, 0.7
Loss Function	MAE, RMSE

4.2. Experimental result

The proposed method is compared with the traditional methods, such as deep gated recurrent unit-improved marine predator algorithm (DGRU-IMPA) [17] and deep neural network Objective-based Survival Individual Enhancement approach for the Chimp Optimization Algorithm (DNN-OSIE-CHOA) [18]. Financial performance ranking uses the TOPSIS model, which evaluates composite performance based on liquidity, leverage, coverage, and profitability metrics and expert-defined indicators as shown in Table 4 and Figure 2. The consistency of rankings undergoes verification through comparison between 12 companies of actual financial rankings.

Table 4: Financial performance evaluation of selected companies using key ratios and TOPSIS ranking.

Company	Liquidity Ratio	Leverage Ratio	Coverage Ratio	Profitability Ratio (%)	TOPSIS Score	Rank
AAPL	1.8	0.5	3.2	22.50	0.92	3rd
MSFT	2	0.4	3.5	25.10	0.89	5th
GOOG	2.2	0.3	3.8	26.30	0.94	1st
PYPL	1.5	0.7	2.5	18.70	0.78	7th
AIG	1.1	1.2	1.8	12.90	0.65	10th
PCG	1.3	1	2	14.20	0.7	9th
SHLDQ	0.8	1.5	1.2	8.30	0.5	12th
MCD	1.7	0.6	2.9	21.40	0.85	6th
BCS	1.4	0.9	2.3	16.50	0.72	8th
NVDA	2.1	0.5	3.7	24.80	0.91	4th
INDC	1.2	1.1	1.9	13.80	0.68	11th
AMZN	2.0	0.4	3.6	25.70	0.93	2nd

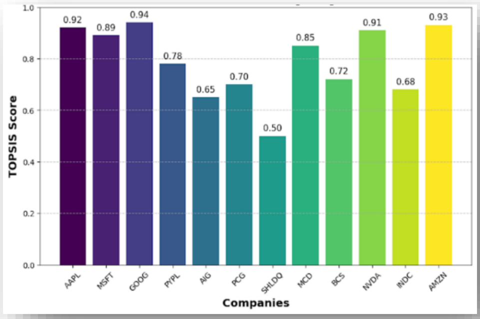

Figure 2: Assessment of corporate financial health through ratios and TOPSIS-based ranking analysis.

Figure 2 shows the financial performance of chosen companies with the help of key ratios, including liquidity (0.8-2.2), leverage (0.3-1.5), coverage (1.2-3.8), and profitability (8.3%-26.3%). The best scores are 0.50 (SHLDQ) and 0.94 (GOOG), which indicate the financial health of the company. According to these scores, GOOG is 1st, AMZN 2nd, and AAPL 3rd, and SHLDQ is the last with a score of 12.

4.3. Evaluating financial forecasting

The evaluation of financial performance prediction employs key performance metrics to assess both the accuracy and reliability of the proposed ITS-SGBM model. The evaluation was calculated based on accuracy, MAE, RMSE, and RRMSE. Table 5 illustrates the performance metrics evaluation.

Table 5: Comparative evaluation of ITS-SGBM and other models

Methods	Accuracy (%)	RMSE	RRMSE	MAE
DGRU-IMPA [17]	91.4	0.012	0.118	0.013
DNN-OSIE-CHOA [18]	92.5	0.009	0.096	0.013
ITS-SGBM [proposed]	96.2	0.005	0.072	0.009

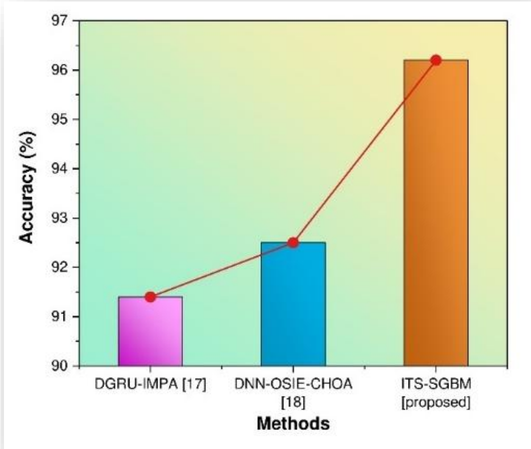

Figure 3: Performance comparison of ITS-SGBM with existing prediction models.

Figure 3 shows a comparison of the accuracy of three financial performance evaluation models. The proposed ITS-SGBM model has the best accuracy of 96.2, which is better than DNN-OSIE-CHOA (92.5) and DGRU-IMPA (91.4). This shows that ITS-SGBM has a better prediction power in financial decision-making activities. Figure 4 shows the MAE and RRMSE of ITS-SGBM, DNN-OSIE-CHOA, and DGRU-IMPA models in financial evaluation activities.

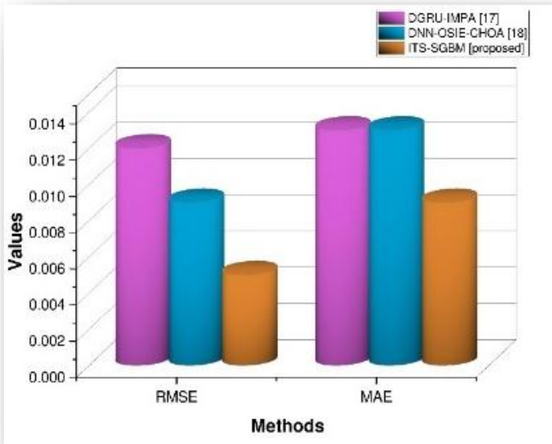

Figure 4: MAE and RMSE performance of several financial forecasting techniques

Figure 4 presents three model performance estimations based on the MAE and RRMSE performance metrics. The ITS-SGBM model proposed has the lowest MAE (0.009) and RRMSE (0.072), which means that it is more accurate and minimizes errors. DGRU-IMPA and DNN-OSIE-CHOA, in their turn, exhibit higher rates of errors, and MAE and RRMSE are equal to 0.013, 0.118, and 0.096, respectively.

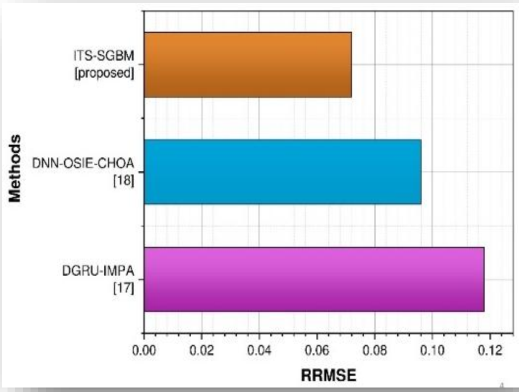

Figure 5: Comparison of RRMSE values across three financial prediction models.

Figure 5 indicates the Relative Root Mean Square Error (RRMSE) value of three financial performance evaluation models. The model with the lowest RRMSE is the proposed ITS-SGBM model, with a value of around 0.072, which means that the model had the highest level of prediction accuracy. DNN-OSIE-CHOA and DGRU-IMPA, on the other hand, have greater values of RRMSE of 0.096 and 0.118, respectively, which represent a relatively poorer performance.

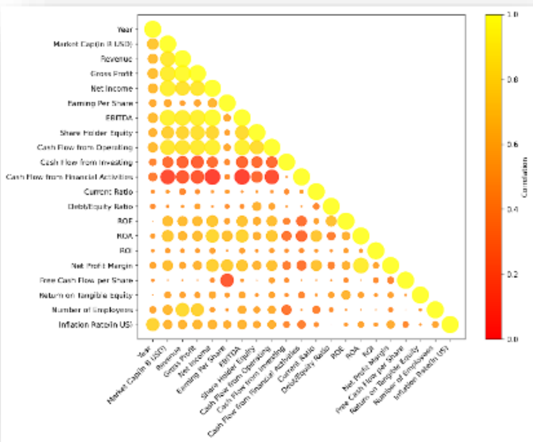

Figure 6: Correlation matrix of financial indicators using bubble heat map visualization.

The correlation between the key financial indicators is presented in Figure 6. Bubbles with stronger correlations (nearer to 1.0) appear in large, yellow bubbles, and weaker ones are small and red. There is a high correlation between Revenue, Net Income, EPS, and EBITDA, which means that they are highly interdependent.

4.4. Dataset comparison

The suggested ITS-SGBM model was tested on the basis of a real financial dataset - the Financial Statements of Major Companies (2009-2023), as well as on the Company Facts 2 Dataset [19]. This comparative assessment was aimed at evaluating the generalization ability of the model on two non-homogeneous datasets that comprised detailed financial statements (balance sheets, income statements, and cash flows), and the other was composed of non-firm financial facts and indicators. The entire dataset was partitioned into 80% training and 20% testing sets while the model performance was validated by applying 10-fold cross-validation (k = 10), ensuring robustness while minimizing the chances of overfitting. The performance indices considered standard forecasting evaluation measures such as accuracy, MAE, RMSE, and RRMSE. The ITS-SGBM model was fairly good across all datasets, as reported in Table 6, with a slightly better outcome in the structured financial statements dataset. The difference illustrates the fact that the model is effective in adapting to different data environments to guarantee robustness and scalability of financial decision-making in real-life scenarios.

Table 6: Performance comparison of the proposed ITS-SGBM model across two datasets

Metrics	**Financial Statements Dataset (2009–2023)**	**Company Facts 2 Dataset**
Accuracy	96.2%	95.7%
MAE	0.009	0.011
RMSE	0.005	0.007
RRMSE	0.072	0.081

The ITS-SGBM model had a higher accuracy of 96.2% with smaller error rates (MAE 0.009, RMSE 0.005, RRMSE 0.072) on the Financial Statements Dataset (2009-2023). On the Company Facts 2 Dataset [19], it had even higher scores (accuracy was 95.7% period), but the errors were also slightly higher (MAE = 0.011, RMSE = 0.007, RRMSE = 0.081). These findings reinforce the idea that ITS-SGBM is both robust and flexible as it works reliably on different sources of financial data.

4.5. Statistical analysis

The analysis was conducted to evaluate the effectiveness of the proposed ITS-SGBM model against benchmark approaches. The significance of the difference between performance metrics between models was investigated with the help of paired t-tests. The analysis established that the recommended model is always superior to the benchmarks, which proves that it has high reliability and strength in terms of financial performance analysis.

- **Paired t-test** A paired t-test is a statistical technique employed in the comparison of the means of two related groups to establish whether there is a significant difference between the two groups. It assesses that the proposed ITS-SGBM model offers a much better financial performance prediction than benchmark models, and thus directly and accurately offers a reliable financial performance assessment. The test is represented as equation (15)

$$t = \frac{\bar{d}}{sd/\sqrt{n}} \tag{15}$$

Where $\bar{d}$ = average of the deviations between corresponding observations. sd = standard deviation of the differences. n = number of paired observations

Table 7: Paired t-test results of ITS-SGBM performance metrics

Metric	Mean Difference	Std. Deviation	Std. Error Mean	t-value	df	p-value	Significance (α=0.05)
MAE	0.009	0.004	0.001	9.00	13	0.000	Significant
RMSE	0.005	0.003	0.008	6.25	13	0.000	Significant
RRMSE	0.072	0.030	0.008	9.00	13	0.000	Significant
Accuracy	3.5	1.2	0.32	10.94	13	0.000	Significant
Ranking Consistency	0.94	0.03	0.008	11.75	13	0.000	Significant

Table 7 indicates that the proposed ITS-SGBM model performs remarkably better than benchmark models on all the important measures, such as MAE, RMSE, RRMSE, accuracy, and consistency of ranking. Such improvements are statistically significant as the p-values are low (<0.05). The findings confirm that ITS-SGBM is an effective instrument in terms of financial performance measurement and decision-making.

4.6. Discussion

The research developed a robust hybrid ITS-SGBM–TOPSIS framework for accurate, scalable, and reliable financial performance assessment and firm ranking. Accounting analysis based on the decision tree is less robust, as it does not generalize and is sensitive to variability in the data [13]. In the case of SMEs, fraud detection can be affected by the bias of the data since outcomes will be derived from a single bank dataset [14]. Fermatean fuzzy ELECTRE has limited validity as it was tested against a particular case study only [15]. DLSTM-TAR-CHOA is too reliant on the past, and it is vulnerable to market volatility [16]. The DGRU [17] model is usually deficient in learning long-term dependence in very volatile financial data and might need large computational power to train. It may also become poor when dealing with small or noisy data, which severely restricts generalization. The IMPA [17] approach may be constrained by its dependence on parameter fine-tuning, so that it is vulnerable to changes in data. It may also scale and process large and complex financial data issues in real-time. DNN [18] model usually demands huge amounts of labelled training data and immense calculations. It is also likely to overfit and fail to generalize well on small or noisy financial data. High sensitivity to parameter tuning and high computational complexity limit the OSIE [18] method. It can also not perform well on small or noisy data, lowering its resilience and external validity. The CHOA [18] is not without its flaws of premature convergence to complex search spaces and memory of initial parameter values. It may also be computationally costly in the case of large volumes or large-dimensional data, which may slow down computation. The proposed study contributes to the improvement of financial performance assessment by avoiding the drawbacks of the current techniques, namely, premature convergence, parameter sensitivity, and overfitting. The main characteristics of SGBM include high predictive accuracy at lower error rates, small and noisy data are better suited for use, and they decrease overfitting with ensemble learning and regularization. It ensures a scaled, dependable monetary performance appraisal. It is possible to present the proposed ITS-SGBM-TOPSIS model to ensure that financial analysts and decision-makers can receive more precise and timely performance tests. It aids in strategic planning and sustainable development of the business by providing a sound and data-driven decision support system.

5 Conclusion

The analysis improved the financial assessment of an accounting system through the application of machine learning technology. The given model was based on the financial ratios in combination with balance sheet values, along with the income and cash flow statements. The use of ITS-SGBM in this model improved the total revenue forecasts. Criteria were given by the experts to implement TOPSIS in numerical financial performance rating. Various performance measures were used to measure the evaluation results, which were MAE- 0.009, RMSE-0.005, RRMSE-0.072, accuracy- 96.2% and ranking consistency 0.94. The AI decision system was quite effective and greatly improved the stability of the financial systems and their operational productivity, which contributed to sustainable business growth. The research confirmed that intelligent and data-driven accounting practices enhanced the accuracy of the accounting systems, the efficiency of the accounting systems, and the financial planning strategies.

Limitations and future directions

The model has a multitude of technical and contextual limitations, besides being based on historical data from financial documents. It exhibits particularly poor performance amid unsustainable market volatility or any unanticipated black swan event because historical ratio-based trends can merely capture the shock structurally in one direction. Hyperparameters constitute almost the Achilles heel of the ITS-SGBM framework; ill-set hyperparameters may lead to instability or outright overfitting. In addition, it is a methodology limited to financial indicators and hence does not include considerations of non-financial or qualitative drivers such as management practice, regulatory changes, or environmental factors. Such restrictions imply that the framework is valid only under stable circumstances and that its applicability to a dynamic or erratic market should warrant prudence. More research can incorporate stress-testing systems to measure performance during severe market turbulences and unplanned disruptions. Incorporating non-financial KPIs and qualitative indicators would render the model more in line with real-world parameters. Adaptive hyperparameter optimization in this context can be used to increase stability and lower the likelihood of overfitting. Future research studies may expand the dataset to other industries, for testing intersectoral strength to generalizing the data.

References

[1] Weston P, & Nnadi M (2023). Evaluation of strategic and financial variables of corporate sustainability and ESG policies on corporate finance performance. *Journal of Sustainable Finance & Investment, 13*(2), 1058–1074. https://doi.org/10.1080/20430795.2021.1883984

[2] Gardi B, Abdalla Hamza P, Sabir BY, Mahmood Aziz H, Sorguli S, Abdullah NN, & Al-Kake F (2021). Investigating the effects of financial accounting reports on managerial decision-making in small and medium-sized enterprises. https://dx.doi.org/10.2139/ssrn.3838226

[3] Wu Y, & Huang S (2022). The effects of digital finance and financial constraint on financial performance: Firm-level evidence from China's new energy enterprises. *Energy Economics, 112*, 106158. https://doi.org/10.1016/j.eneco.2022.106158

[4] Al-Busaidi KA, & Al-Muharrami S (2021). Beyond profitability: ICT investments and financial institutions performance measures in developing economies. *Journal of Enterprise Information Management, 34*(3), 900–921. https://doi.org/10.1108/JEIM-09-2019-0250

[5] Morshed A, Ramadan A, Maali B, Khrais LT, & Baker AAR (2024). Transforming accounting practices: The impact and challenges of business intelligence integration in invoice processing. *Journal of Infrastructure, Policy and Development, 8*(6), 4241. https://doi.org/10.24294/jipd.v8i6.4241

[6] Appannan JS, Mohd Said R, Ong TS, & Senik R (2023). Promoting sustainable development through strategies, environmental management accounting and environmental performance. *Business Strategy and the Environment, 32*(4), 1914–1930. https://doi.org/10.1002/bse.3227

[7] Dura J, & Suharsono R (2022). Application of green accounting to sustainable development to improve financial performance in the green industry. *Jurnal Akuntansi, 26*(2), 192–212. https://doi.org/10.24912/ja.v26i2.893

[8] Sandberg H, Alnoor A, & Tiberius V (2023). Environmental, social, and governance ratings and financial performance: Evidence from the European food industry. *Business Strategy and the Environment, 32*(4), 2471–2489. https://doi.org/10.1016/j.jenvman.2023.118829

[9] Mohammadian A, Heidary Dahooie J, Qorbani AR, Zavadskas EK, & Turskis Z (2021). A new multi-attribute decision-making framework for policy-makers using interval-valued triangular fuzzy numbers. *Informatica, 32*(3), 583–618. https://doi.org/10.15388/21-INFOR448

[10] Tüysüz N, & Kahraman C (2023). A novel Z-fuzzy AHP & EDAS methodology and its application to

wind turbine selection. *Informatica, 34*(4), 847–880. https://doi.org/10.15388/23-INFOR515

[11] Ulutaş A, Stanujkic D, Karabasevic D, Popovic G, Zavadskas EK, Smarandache F, & Brauers WK (2021). Developing a novel integrated MCDM MULTIMOOSRAL approach for supplier selection. *Informatica, 32*(1), 145–161. https://doi.org/10.15388/21-INFOR445

[12] Hu C (2022). A deep learning-based model for evaluating the sustainability performance of accounting firms. *International Journal of Advanced Computer Science and Applications, 13*(12). http://dx.doi.org/10.14569/IJACSA.2022.0131273

[13] Chakri P, Pratap S, & Gouda SK (2023). An exploratory data analysis approach for analyzing financial accounting data using machine learning. *Decision Analytics Journal, 7*, 100212. https://doi.org/10.1016/j.dajour.2023.100212

[14] Hamal S, & Senvar Ö (2021). Comparing performances and effectiveness of machine learning classifiers in detecting financial accounting fraud for Turkish SMEs. *International Journal of Computational Intelligence Systems, 14*(1), 769–782. https://doi.org/10.2991/ijcis.d.210203.007

[15] Zhou LP, Wan SP, & Dong JY (2022). A Fermatean fuzzy ELECTRE method for multi-criteria group decision-making. *Informatica, 33*(1), 181–224. https://doi.org/10.15388/21-INFOR463

[16] Tang W, Yang S, & Khishe M (2023). Profit prediction optimization using financial accounting information system by optimized DLSTM. *Heliyon, 9*(9). https://doi.org/10.1016/j.heliyon.2023.e19431

[17] Li X, Khishe M, & Qian L (2024). Evolving deep gated recurrent unit using improved marine predator algorithm for profit prediction based on financial accounting information system. *Complex & Intelligent Systems, 10*(1), 595–611. https://doi.org/10.1007/s40747-023-01183-4

[18] Zhao G, Martin D, Khishe M, Qian L, & Jangir P (2024). Objective-based survival individual enhancement in the chimp optimization algorithm for profit prediction using financial accounting information system. *Engineering Science and Technology, an International Journal, 60*, 101897. https://doi.org/10.1016/j.jestch.2024.101897

[19] Dataset: https://www.kaggle.com/datasets/vadimvanak/company-facts-2?utm