# Enhancing Financial Forecasting for Small Businesses: A Robust Approach to Revenue and Expense Prediction

**Safaa Zaman[1,*], El-Sayed M. El-Kenawy[2,3]**

[1]Information sciences department, College of life sciences, Kuwait University, Kuwait
[2]Department of Communications and Electronics, Delta Higher Institute of Engineering and Technology, Mansoura 35111, Egypt
[3]Jadara Research Center, Jadara University, Irbid 21110, Jordan

Emails: S.3zaman@ku.edu.kw; skenawy@ieee.org

**Abstract**

This study addresses the critical challenge of financial forecasting for small businesses, which often struggle with fluctuating demand, seasonal sales patterns, and tight profit margins. Accurate forecasting is essential for optimizing resources, improving profitability, and making data-driven decisions in a dynamic market. To enhance the accuracy and efficiency of forecasting models, this paper introduces a novel approach combining machine learning models with metaheuristic optimization algorithms. Specifically, the Dynamic Attention Recurrent (DAR) model optimized with Logarithmic Transformation (LogTrans) is evaluated at various stages. In the baseline evaluation, the DAR + LogTrans model demonstrated outstanding performance with an MSE of 0.00075, RMSE of 0.0274, and R-squared of 0.861, indicating its strong predictive capability. After applying optimization techniques, DAR + LogTrans achieved remarkable improvements, reaching an MSE of 1.88E-07, RMSE of 4.36E-04, and R-squared of 0.968, showcasing substantial gains in accuracy and generalization. The results emphasize the potential of metaheuristic optimization, such as the Whale Optimization Algorithm (WAO), Bat Algorithm (BA), and Particle Swarm Optimization (PSO), in improving model performance. These findings provide valuable insights for small business owners seeking to implement advanced forecasting models that can adapt to market fluctuations. The optimized models, particularly DAR + LogTrans, offer a powerful tool for improving decision-making, managing cash flow, and enhancing operational efficiency, with significant implications for the future of financial forecasting in small businesses.

**Keywords:** Financial Forecasting; Small Business Analytics; Revenue Prediction; Expense Management; Economic Decision-Making

## 1 Introduction

The role of data science in modern business analytics has witnessed an exponential rise, particularly for small businesses operating in competitive markets, such as coffee shops [1], [2], [3], [4]. These businesses often face a range of operational challenges, from fluctuating customer demand and seasonal sales patterns to tight profit margins [5], [6]. As these businesses work to remain competitive, the ability to efficiently manage finances becomes a critical factor in ensuring survival and fostering growth. Financial management in such businesses goes beyond just tracking sales and expenses; it involves strategic decision-making to optimize resources, reduce inefficiencies, and improve profitability. Over the last decade, the increased use of technology has allowed business owners to gain real-time insights into their financial performance, which has become an essential tool in running day-to-day operations. As a result, small businesses increasingly rely on accurate

DOI: https://doi.org/10.54216/JSDGT.050201

financial tracking and forecasting methods to gain a competitive edge, plan for future growth, and remain resilient in volatile economic conditions [7], [8].

To achieve financial sustainability and growth, small businesses must adopt tools that provide visibility into their financial health. Financial dashboards, revenue forecasting tools, and expense management systems are now integral to daily business operations [9], [10]. These tools help businesses understand their sales patterns, manage their expenses, track profit margins, and identify trends. This level of financial visibility enables business owners to make more informed decisions regarding inventory, pricing strategies, and staffing. Furthermore, by forecasting future revenues and expenses, business owners are able to anticipate and plan for potential challenges, ensuring that they can adapt to changes in consumer behavior, market conditions, and external economic factors. In essence, these financial tools empower small businesses to not only keep their operations running smoothly but to continuously improve their financial strategies, ensuring long-term success and stability.

Financial forecasting is essential to the survival and growth of businesses, particularly those operating in industries where demand is not consistent year-round. In sectors like the coffee shop industry, where customer footfall and sales can vary significantly based on the season or even local events, accurate financial forecasts become crucial [11], [12]. Financial forecasting enables small businesses to manage their cash flow effectively, prepare for upcoming expenses, and identify any financial shortfalls before they become critical [13], [14]. By predicting revenue and expenses, businesses can optimize their inventory management, adjust their pricing strategies to align with consumer trends, schedule their staff based on expected sales, and tailor their marketing efforts to generate higher returns. More importantly, accurate financial forecasts provide businesses with the agility they need to quickly adjust to changing circumstances, whether it's a change in consumer preferences or an unexpected supply chain disruption. For small businesses, this adaptability is often the key to remaining competitive in a market that is constantly evolving. In industries such as hospitality, where margins are typically thin and external factors can have a direct impact on profitability, financial forecasting offers a means to make informed decisions that improve business sustainability and long-term growth.

Machine learning (ML) has proven to be an invaluable tool in transforming traditional financial forecasting methods. While traditional statistical methods have been widely used in financial forecasting for decades, they often struggle to handle the complexity and non-linearity present in modern financial datasets. Machine learning, on the other hand, is well-equipped to handle large and complex datasets, uncovering hidden patterns and relationships that would be difficult for traditional models to detect. By using algorithms that learn from historical data, ML models are able to generate highly accurate predictions, which are crucial for businesses trying to anticipate future revenues, expenses, and profits. These machine learning models, particularly those designed for time-series analysis, have shown great promise in handling the inherent temporal dependencies present in financial data. For instance, machine learning techniques such as neural networks, decision trees, and ensemble methods can identify trends, detect seasonality, and provide deeper insights into consumer behavior. This paper aims to apply various machine learning models to predict revenue and expenses for a small-town coffee shop based on historical financial data, demonstrating how these models can outperform traditional methods and provide small businesses with more reliable and actionable forecasts.

While financial datasets offer great potential for making accurate predictions, they also present a range of challenges that must be addressed to ensure the effectiveness of predictive models. One of the primary challenges in financial data is the issue of high dimensionality. Financial datasets, particularly those from small businesses like coffee shops, can include a wide variety of features, including daily sales data, payroll information, vendor charges, and operating expenses. Each of these features represents a different aspect of business operations and provides valuable insights into the financial health of the business. However, when multiple features are included in a predictive model, the complexity of the data increases significantly. High-dimensional datasets introduce challenges for traditional statistical models and machine learning algorithms, making them prone to overfitting, where the model becomes too complex and loses its ability to generalize to new data. The computational burden also increases with more features, which can lead to longer training times and higher computational costs. Managing and analyzing high-dimensional data, therefore, requires sophisticated techniques to reduce dimensionality while preserving critical information that contributes to accurate forecasting. Feature selection and dimensionality reduction techniques are essential for improving the efficiency and accuracy of financial forecasting models.

In addition to high dimensionality, feature redundancy is another challenge in financial prediction models. Many financial features tend to be highly correlated with one another. For example, sales data from different product categories such as hot drinks and pastries may exhibit similar trends, leading to redundancy in the dataset. Redundant features convey similar information to the model, which can result in multicollinearity. Multicollinearity can destabilize regression coefficients and make it difficult to interpret the impact of individual features on the model's predictions. It can also lead to inefficiencies in the model, as unnecessary features can increase the model's complexity without providing additional valuable information. Identifying and removing redundant features is crucial to improving the stability and interpretability of predictive models. Feature selection techniques, such as correlation analysis and principal component analysis (PCA), are often employed to address feature redundancy and reduce the number of features in the dataset.

Hyperparameter tuning is another critical challenge in machine learning-based financial forecasting models. Hyperparameters, such as the learning rate, regularization coefficients, and the number of layers in neural networks, play a significant role in the performance of the model. Selecting the optimal set of hyperparameters can significantly affect the model's predictive accuracy and generalization ability. However, the process of hyperparameter optimization is often time-consuming and requires extensive experimentation. The improper selection of hyperparameters can lead to suboptimal performance, with the model either underfitting or overfitting the data. Underfitting occurs when the model is too simple to capture the complex patterns in the data, while overfitting happens when the model becomes too complex and captures noise instead of the true underlying trends. Efficient hyperparameter optimization techniques, such as grid search, random search, and more advanced metaheuristic algorithms, are necessary to fine-tune the model and achieve the best possible performance.

Overfitting is a common problem in predictive models, particularly in financial forecasting. Overfitting occurs when the model fits too closely to the training data, capturing not only the underlying patterns but also noise and outliers that do not generalize well to new data. As a result, the model's performance on unseen data suffers. In financial forecasting, where the data is dynamic and subject to external factors such as market changes and consumer behavior, overfitting can significantly reduce the model's reliability. To mitigate overfitting, techniques such as regularization (which adds a penalty for complex models), cross-validation (which evaluates the model's performance on multiple data subsets), and early stopping (which halts training when the model's performance stops improving) are commonly employed. These methods help ensure that the model generalizes well to new, unseen data and performs reliably in real-world financial forecasting applications.

The primary objective of this study is to develop and compare several machine learning models for forecasting revenue and expenses based on the financial data from the coffee shop. This study will evaluate a variety of machine learning models, with a particular focus on those capable of handling time-series data, which is a common characteristic of financial forecasting tasks. Time-series models are designed to capture temporal dependencies in the data, making them ideal for forecasting future financial trends based on past behavior. The study will assess the models' ability to identify and learn from underlying patterns, trends, and seasonality in the data, which are essential for generating accurate revenue and expense forecasts for a small business.

In addition to the development of predictive models, the study aims to optimize these models using metaheuristic algorithms for feature selection and hyperparameter optimization. Metaheuristic algorithms, such as Particle Swarm Optimization (PSO) and Whale Optimization Algorithm (WOA), are highly effective for solving complex optimization problems, including selecting the most relevant features and fine-tuning hyperparameters. These techniques will be applied to improve the models' accuracy and computational efficiency. By refining the models' inputs and settings, metaheuristic algorithms can enhance predictive accuracy while reducing the computational cost associated with training and evaluation. The integration of feature selection and hyperparameter optimization is expected to improve the robustness of the models, ensuring that they provide reliable financial forecasts for small businesses, even when dealing with complex, high-dimensional data.

The study also seeks to demonstrate the performance and efficiency gains achieved through the optimization process. By applying feature selection and hyperparameter optimization techniques, the models are expected to show significant improvements in predictive accuracy, reduced error rates, and faster computation times. These improvements will not only enhance the quality of financial forecasts but will also make the models more practical and scalable for real-time forecasting in small businesses. The ability to provide real-time financial predictions will enable small businesses to make more informed decisions quickly, improving their operational agility and helping them remain competitive in dynamic market conditions.

This study introduces a novel methodology that combines machine learning models with metaheuristic optimization algorithms to improve financial forecasting for small businesses. The hybrid approach leverages the strengths of machine learning in detecting complex patterns and the power of metaheuristic algorithms in optimizing model inputs and parameters. By combining these two techniques, the study aims to address common challenges in financial prediction, such as high-dimensionality, feature redundancy, and hyperparameter sensitivity, thereby improving the accuracy and efficiency of financial forecasts.

The paper will provide a detailed comparative analysis of machine learning models before and after optimization, showcasing the impact of metaheuristic optimization on model performance. The models will be evaluated using performance metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared ($R^2$), which will highlight improvements in predictive accuracy and computational efficiency. The comparison will demonstrate the effectiveness of applying metaheuristic optimization techniques to enhance the performance of financial forecasting models.

The integration of feature selection and hyperparameter optimization will be explored in depth, highlighting how these two techniques complement each other to improve model performance. By applying both optimization techniques simultaneously, the research aims to show how this integrated approach can lead to better forecasting results, ultimately improving decision-making and operational efficiency for small businesses.

### 1.1 Structure of the Paper

The paper is organized as follows:

- **Section 2**: Provides a detailed description of the dataset and the preprocessing methods used to prepare the data for analysis.
- **Section 3**: Discusses the machine learning models and metaheuristic algorithms employed in the study, explaining their suitability for financial forecasting tasks.
- **Section 4**: Presents the empirical results, including a comparison of baseline models and optimized models.
- **Section 5**: Concludes the paper and outlines directions for future research.

## 2 Literature Review

The integration of Artificial Intelligence (AI) into industries and businesses has emerged as a crucial factor in driving innovation, particularly with significant advancements in recent years. The adoption of AI in various sectors such as retail, healthcare, and telecommunications has revolutionized operations, improving efficiency, enhancing customer experience, and generating new market opportunities. Despite AI's potential, certain industries remain skeptical of its practical application, raising concerns over the cost and implementation challenges. While tech giants like Facebook, Google, and Baidu have made considerable investments in AI, with an estimated 2 trillion globally spent on AI research and development in 2021 alone, many traditional industries are still grappling with how to allocate funds toward AI initiatives. Despite these challenges, AI-driven technologies, particularly in self-driving cars, deep learning, and robotics, continue to gain momentum. These areas promise to deliver significant benefits to businesses, enabling them to gain real-world competitive edges as they prepare for digital changes. AI, particularly deep learning, continues to be a key area of focus, contributing to growth in various sectors, including autonomous driving and AI-driven marketing applications [15].

In the context of financial investments, Environmental, Social, and Governance (ESG) factors have become a focal point, especially in the wake of the COVID-19 pandemic. The pandemic heightened the importance of sustainability in investment decisions, leading to a growing emphasis on ESG metrics as a critical determinant of investment performance. Studies indicate that funds adhering to sustainable practices often yield better returns compared to those that do not. However, research exploring the application of AI algorithms to analyze ESG data has been limited, primarily due to challenges in gathering sufficient ESG-related datasets. One study addresses this gap by proposing an AI-based approach to ESG data analysis, particularly focusing on governance and social datasets. By leveraging Natural Language Processing (NLP) algorithms, the study introduced a model that can predict a firm's ESG rankings. Furthermore, the study investigates adversarial attacks on ESG datasets, highlighting the importance of robustness in AI models. The results demonstrated the effectiveness of AI in ESG analysis, offering potential benefits for small businesses seeking to assess their ESG performance more efficiently. This work provides a clear path for integrating AI into ESG data processing and prediction, which is crucial for enhancing transparency and decision-making in the financial sector [16].

Financial forecasting, particularly the prediction of corporate bankruptcy, is a significant area of concern for both researchers and practitioners in the financial sector. Accurate bankruptcy prediction models are critical for assessing financial health, managing risks, and making informed investment decisions. A comprehensive study compared several machine learning models, including Deep Neural Networks (DNN), Random Forest (RF), and Support Vector Machine (SVM), to predict bankruptcy in Tunisian companies. The dataset used includes 25 financial ratios from 732 companies between 2011 and 2017. The findings suggest that DNN outperforms other models in terms of accuracy, highlighting its superior predictive capabilities. In contrast, Random Forest performed well in terms of interpretability and robustness. This research emphasizes the growing role of machine learning in risk management and underscores the importance of adopting AI techniques for early detection of financial distress. It illustrates how machine learning models can enhance financial forecasting and improve decision-making by providing more accurate predictions of bankruptcy risk [17].

The stock market remains one of the most difficult fields to predict, with a high level of volatility and uncertainty. The study investigated the application of Machine Learning (ML), Deep Learning (DL), and Deep Reinforcement Learning (DRL) in Quantitative Finance (QF) and stock market predictions. DRL has gained significant attention in recent years for its ability to combine price prediction and trading signal generation into a cohesive system, enabling the construction of automated trading strategies. The study outlines how DRL agents can optimize trading performance in real-time by adjusting strategies based on market conditions. This paper not only highlights the potential of DRL in algorithmic trading but also addresses the challenges associated with applying these techniques, such as the need for extensive data and the complexity of tuning hyperparameters. The authors propose future research directions to further explore DRL's applications in financial markets and improve its scalability and adaptability in live trading environments [18].

The accuracy of financial failure prediction is crucial for companies to mitigate risks and adapt to changes in the financial environment. Traditional financial forecasting models often struggle with imbalanced datasets. One study compared deep learning techniques, specifically Long Short-Term Memory (LSTM) networks and Multilayer Perceptron (MLP), with traditional ensemble methods such as Random Forest (RF), Support Vector Machine (SVM), and K-Nearest Neighbor (KNN). The study used datasets from Spanish, Taiwanese, and Polish companies and demonstrated that LSTM networks outperformed traditional classifiers in terms of accuracy and recall. The integration of oversampling techniques such as SMOTE-ENN further improved model performance, reducing the Type II error rate and improving predictive accuracy. This paper emphasizes the power of deep learning to improve the predictive accuracy of financial distress and failure prediction models, offering a promising solution to the limitations of traditional forecasting approaches [19].

E-commerce companies face unique financial risks due to the volatile nature of the industry. Bankruptcy and liquidation rates have been rising, and early risk detection is essential for preventing financial crises. One research focused on the development of an AI-driven early warning system for financial risks in e-commerce businesses. By using deep learning algorithms, the proposed model is able to detect financial crises early and provide companies with the necessary time to take corrective actions. The paper outlines how the integration of financial security systems with AI can help e-commerce companies allocate resources more effectively and avoid financial failure. This research highlights the importance of AI in managing financial risks in fast-paced industries and provides valuable insights into how businesses can stay competitive in a rapidly changing economic landscape [20].

Financial management evaluation models have also benefited from the incorporation of deep learning techniques. One study introduced a financial management evaluation model based on deep belief networks (DBN) and compared it with traditional models such as KNN, SVM-RBF, and SVM linear. The results showed that the deep belief network model outperformed other classical deep learning methods in terms of accuracy. This paper underscores the potential of deep learning in improving financial management practices by providing more accurate predictions and helping businesses identify risks earlier in the decision-making process. The study also demonstrated how DBN can handle complex financial datasets and offer a more efficient and robust evaluation framework for financial management [21].

Fraud detection is an ongoing challenge in financial institutions, as traditional rule-based methods are often insufficient to detect increasingly sophisticated fraudulent activities. One paper reviewed recent advancements in machine learning and deep learning algorithms for fraud detection. The paper categorized these algorithms based on their learning techniques, such as supervised, unsupervised, and reinforcement learning. The study emphasized the importance of feature engineering and data pre-processing to improve the effectiveness of fraud detection systems. Additionally, the paper discussed key challenges such as imbalanced datasets, adversarial attacks, and model interpretability, which can affect the accuracy and performance of fraud detection systems. This review provides valuable insights into the state-of-the-art methods in fraud detection and paves the way for future research to address the challenges in this critical area [22].

Machine learning (ML) and deep learning (DL) technologies are integral to the transformation of various industries. One study provides a comprehensive view of ML and DL algorithms and their applications in industries such as healthcare, finance, agriculture, and transportation. It analyzes the effectiveness of these technologies in addressing real-world problems, with a focus on supervised, unsupervised, semi-supervised, and reinforcement learning techniques. The paper also discusses emerging trends in ML and DL, such as hybrid models and generative AI, and highlights the potential for further development in these areas. This study offers a roadmap for researchers and practitioners looking to apply AI-driven solutions in diverse industries, helping to advance the integration of AI technologies in real-world applications [23].

The use of Enterprise Resource Planning (ERP) systems in financial management has become increasingly prevalent in modern businesses. One study focused on the development of a financial risk prediction model based on deep learning, specifically using a Temporal Convolutional Network-Long Short-Term Memory (TCN-LSTM) structure. The research introduces an optimization algorithm to improve the model's accuracy and stability. The results demonstrate the superiority of the deep learning-based risk prediction model compared to traditional benchmark models, showcasing the potential of AI-driven financial management solutions to help companies allocate resources efficiently and enhance their financial forecasting capabilities [24].

Credit risk prediction and financial distress prediction are essential components of financial decision-making. One study proposes a new financial distress prediction model that combines the Adaptive Whale Optimization Algorithm (AWOA) with deep learning techniques. The AWOA is used to optimize the hyperparameters of a deep neural network (DNN) model, significantly improving the accuracy of the prediction. The study shows that the AWOA-DL model outperforms other compared techniques, achieving an average accuracy of 95.8%. This research demonstrates the potential of combining optimization algorithms with deep learning to improve the forecasting capabilities of financial institutions and enhance credit risk assessments [25].

The application of machine learning and deep learning in credit risk assessment has gained momentum, especially in emerging markets where individuals often lack access to traditional forms of collateral or identification. One study explores how AI can address the challenges of information asymmetry and moral hazard in credit risk assessment. By utilizing alternative data sources, such as public data, AI-driven models can enhance the accuracy of credit risk evaluations and provide underserved populations with better access to credit. The study advocates for greater investment in AI technologies by financial institutions to improve credit risk analysis and promote financial inclusion in emerging markets [26].

The combination of Big Data, Internet of Things (IoT), and cloud computing technologies has created new opportunities for financial forecasting, particularly in predicting the risk of financial crises in small and medium-sized enterprises (SMEs). One study develops an oppositional ant lion optimizer-based feature selection (OALOFS) with machine learning-enabled classification (MLC) for financial crisis prediction (FCP). The OALOFS-MLC model uses Hadoop MapReduce for big data management and achieves improved classification results compared to recent approaches. This study demonstrates the potential of combining advanced optimization algorithms with machine learning techniques to enhance financial forecasting in the context of big data environments [27].

Predicting bankruptcies and assessing credit risk are essential issues in the finance sector. One research focuses on the use of an Adaptive Whale Optimization Algorithm (AWOA) in conjunction with deep learning (DL) for financial distress prediction. The AWOA-DL model optimizes the hyperparameters of a deep neural network

(DNN), achieving an impressive accuracy rate of 95.8%. This study highlights the importance of optimization algorithms in enhancing the performance of deep learning models and improving the accuracy of financial distress prediction [28].

Finally, one study explores the relationship between financial indicators, R&D investment, and the educational background of entrepreneurs in the context of corporate financial forecasting. By incorporating recurrent neural networks (RNNs) into the financial prediction model, the study demonstrates how these networks can improve forecasting performance, especially for small and medium-sized enterprises (SMEs). The paper also shows that the number of hidden layers and the memory function dimensions of the BP model have a significant impact on the accuracy of financial predictions [29].

## 3 Materials and Methods

### 3.1 Dataset Description

The dataset used in this study simulates the financial records of a small-town coffee shop for the period from January 2022 to December 2023. This dataset was created to reflect the financial activities of a typical small business operating in the hospitality sector, specifically a coffee shop that experiences fluctuating demand throughout the year due to seasonal variations. For example, demand tends to peak during colder months or around holidays, while sales tend to dip during warmer months or slower periods of the year. This seasonality is crucial for understanding the business's financial trends, which is why the dataset includes comprehensive information for a full two-year period. The dataset is composed of five CSV files, each containing different types of financial data, such as daily sales deposits, monthly transfers between accounts, weekly expenses, and payroll information. These files are structured to provide a comprehensive overview of the coffee shop's financial operations and to help predict future sales and expenses more effectively.

The dataset's structure includes various types of data, including income from sales, expenses from operations, payroll data for employees, and various financial transfers between accounts. These diverse data points allow for the analysis of multiple aspects of the coffee shop's operations. The financial records in the dataset provide a valuable resource for exploring how external factors, such as local events or seasonal changes, affect revenue and expenditure. The data also provides insight into operational efficiency and helps in identifying opportunities for cost savings. By analyzing this dataset, it is possible to predict revenue fluctuations, optimize operational costs, and better manage the business's cash flow.

The coffee shop dataset features several key elements: sales deposits from hot drinks, cold drinks, pastries, and sandwiches; operating expenses including utilities, rent, supplies, and other miscellaneous costs; and payroll data for employees, including wages, tips, and tax deductions. The monthly transfer files allow for tracking inter-account movements of funds, providing insight into the business's overall cash flow management. By analyzing these elements, we gain a comprehensive understanding of the business's financial health.

The coffee shop dataset was specifically designed to represent the financial dynamics of a small business in a competitive market. It includes essential features such as transaction descriptions, amounts, payroll details, tax deductions, and balances across various accounts such as checking accounts, credit card accounts, and payroll accounts. These features enable comprehensive financial analysis, helping to predict trends, assess profitability, and evaluate the sustainability of the business over time.

The primary target variable for prediction in this study is the coffee shop's daily, monthly, or quarterly revenue. This revenue is influenced by a variety of factors, including the sales of specific product categories such as hot drinks, pastries, and sandwiches. These product categories often show differing seasonal patterns that impact the overall revenue, which must be factored into any predictive model. A secondary target of the model is forecasting expenses related to payroll, utilities, and supplies. These expenses, along with the predicted revenue, are crucial for understanding the coffee shop's profitability, cost structure, and cash flow.

The dataset does not incorporate external data sources such as regional economic indicators or demographic information. While such external data could enhance the robustness of the forecasting model by providing a more comprehensive view of the market environment, the current study focuses solely on the internal financial data of the coffee shop. Nonetheless, future work could integrate external data sources to improve model predictions and better account for factors such as changes in local economic conditions, consumer behavior, and competitive dynamics in the local market.

The dataset has been divided into separate subsets for training, validation, and testing purposes. The training set will be used to fit the models, while the validation set will be employed to tune hyperparameters and assess model performance during the development process. Finally, the test set will be used to evaluate the final model's performance on unseen data, ensuring that the model generalizes well and is not overfitting to the training data. By separating the dataset into these subsets, the study aims to ensure that the models developed can provide reliable predictions for future financial performance, even in new or changing market conditions. In this analysis, we explore the distribution and variation of salary, net pay, tips, and tax by role within a company. The figure below presents several visualizations that highlight the differences across various employee roles such as Owner, Barista, Manager, Independent Contractor, and Seasonal Barista. The first plot (top left) shows the **salary density by role** using a Kernel Density Estimate (KDE) to visualize how gross pay is distributed across different roles. The second plot (top right) presents the **net pay distribution** for each role using a swarm plot to show individual data points. The third plot (bottom left) illustrates the **tips breakdown by role** with a violin plot, emphasizing the role of tips in employee compensation. Finally, the **tax variation by role** is shown using a box plot in the bottom right, indicating the distribution of federal tax values for each role (see Figure 1).

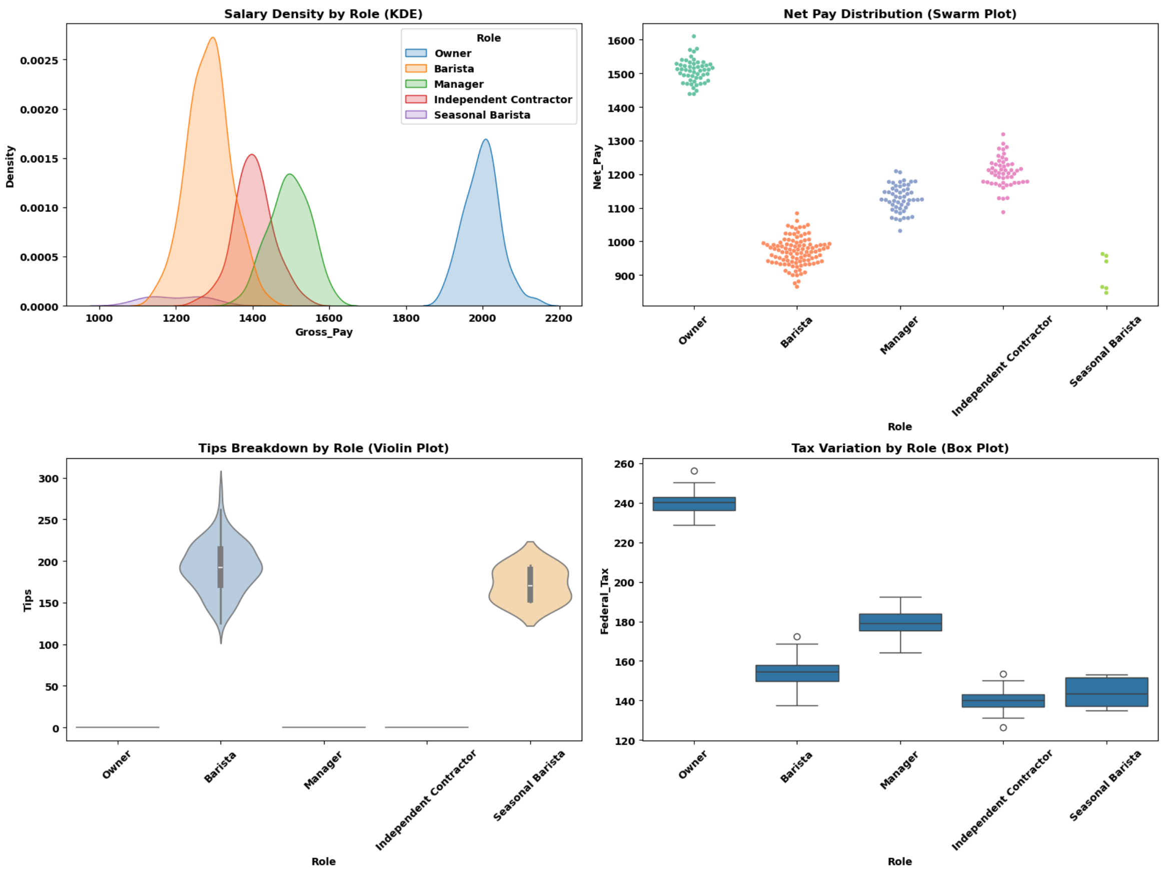


Figure 1: Salary, Net Pay, Tips, and Tax Variation by Role

The figure below depicts the relationship between **gross pay** and **net pay** across different roles in the company. The scatter plot visualizes how net pay increases as gross pay rises, with distinct color coding for various roles including Owner, Barista, Manager, Independent Contractor, and Seasonal Barista. Each point represents an individual role, and the spread highlights how the relationship varies for different job positions. This plot provides insight into how each role's net pay scales relative to their gross pay, with Owners having the highest net pay and Seasonal Baristas showing a more moderate increase (see Figure 2).

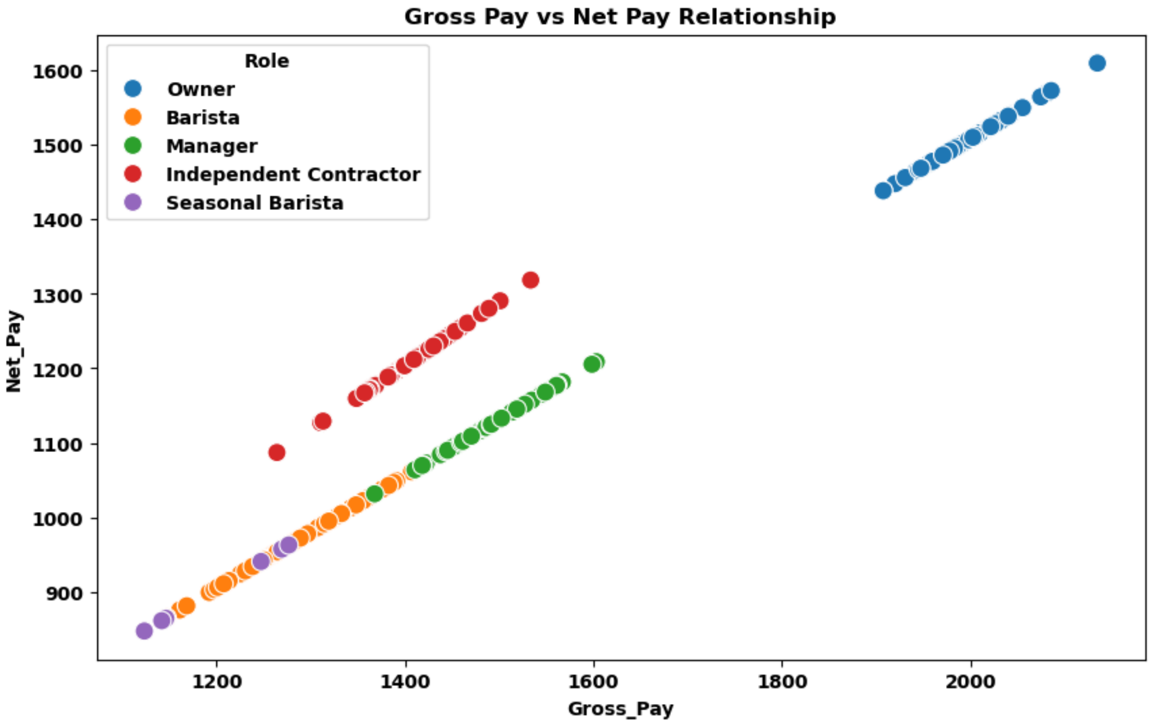


Figure 2: Gross Pay vs Net Pay Relationship

The figure below illustrates the monthly payroll expenses over the course of 2022 and 2023. The line plot shows the variation in payroll expenses for each month, with notable spikes in July 2022. These fluctuations may be due to several factors, such as additional bonuses, seasonal pay variations, or other payroll adjustments that occur periodically. Understanding these variations is essential for budgeting and financial planning within the company. The plot highlights the general trend and provides insight into the consistency or volatility of payroll costs over time (see Figure 3).

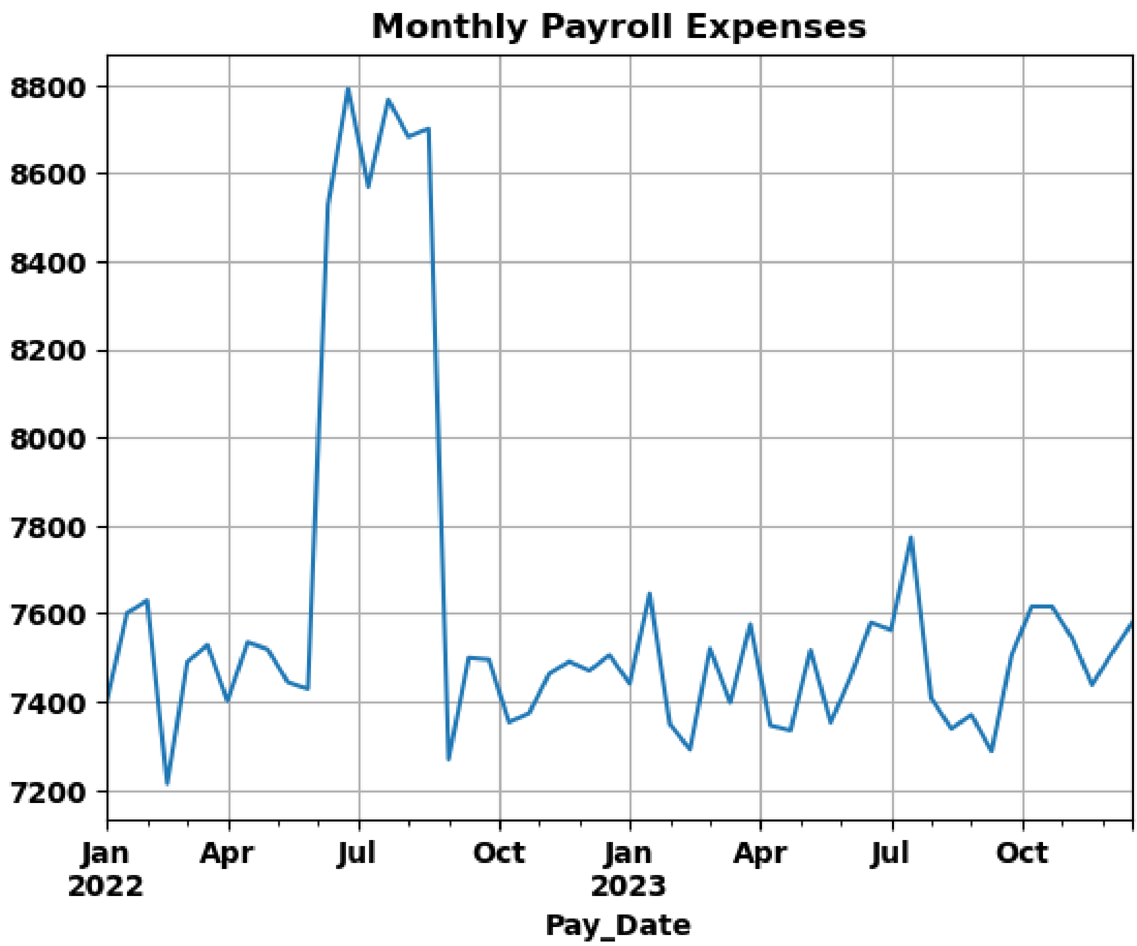


Figure 3: Monthly Payroll Expenses

### 3.2 Data Preprocessing

The first step in preparing the dataset for machine learning modeling is data preprocessing. Given the financial nature of the dataset, handling missing data, encoding categorical variables, scaling numerical features, and analyzing feature correlations are critical steps to ensure that the models perform optimally.

Missing data is a common issue in financial datasets, and it can occur for various reasons, such as incomplete transactions or missing records. To address this, imputation techniques are applied to fill in missing values. For numerical features, the missing values are imputed using either the mean or median of the corresponding feature. The median is typically used when the data distribution is skewed or contains outliers, as it is less sensitive to extreme values than the mean. This imputation ensures that the data remains complete and usable for training the models, preventing any data points from being excluded unnecessarily. For categorical variables, missing values are filled using the most frequent category (mode), ensuring that the encoding is not biased by the missing entries. By using these imputation methods, the integrity of the dataset is maintained, and the models can be trained effectively without the risk of missing data skewing results.

Another critical aspect of data preprocessing is the encoding of categorical features. The dataset includes several categorical variables, such as transaction descriptions (e.g., "hot drinks," "cold drinks," "pastries") and employee roles (e.g., "Owner," "Barista," "Manager"). Machine learning models, particularly those based on linear regression or neural networks, require numerical input, and therefore, categorical features must be converted into a numerical format. For non-ordinal categorical variables, such as transaction descriptions, one-hot encoding is applied. One-hot encoding creates a binary column for each category, indicating the presence or absence of that category for a given observation. This encoding ensures that the model treats each category as a separate, non-ordinal feature. On the other hand, for ordinal features like employee roles, where there is a natural ranking (e.g., "Barista" < "Manager" < "Owner"), label encoding is used, where each category is assigned a numerical value. This approach preserves the inherent order among the categories.

Feature scaling is another essential step in data preprocessing. Some machine learning algorithms, such as those based on gradient descent, are sensitive to the scale of the input features. Features like transaction amounts, which can have a wide range (e.g., a small sale of $5versusalargesaleof500$), need to be standardized. Standardization involves transforming the data such that the mean of each feature is 0 and the standard deviation is 1. This process prevents features with larger values from dominating the learning process and helps gradient-based algorithms converge more quickly. For example, without scaling, a feature representing large transaction amounts might disproportionately influence the model's decisions, leading to biased predictions. Standardizing the features ensures that each variable contributes equally to the model's predictions.

Feature correlation analysis is also an important preprocessing step, particularly in financial datasets where certain features may be highly correlated with each other. For example, sales from hot drinks and pastries may show a strong correlation because they are often purchased together. Highly correlated features can lead to multicollinearity, where the model cannot distinguish the individual contributions of each correlated feature, leading to inefficiency and poor model interpretability. To identify these correlations, Pearson correlation coefficients are calculated between features. Features that exhibit a high correlation (above a predefined threshold, such as 0.8) are flagged for potential removal or transformation. By eliminating redundant features, the model becomes more interpretable, and the risk of overfitting is reduced. Additionally, removing correlated features can help to improve the model's performance by reducing noise and unnecessary complexity in the data.

In summary, the data preprocessing steps involve handling missing values, encoding categorical variables, scaling numerical features, and analyzing feature correlations. These preprocessing techniques are essential to ensure that the data is in an optimal form for machine learning models and that the models can produce reliable and accurate predictions. By carefully preparing the data, the study aims to ensure that the financial forecasting models are both efficient and effective, providing valuable insights for the coffee shop's financial management.

### 3.3 Deep Learning Models

#### 3.3.1 Model Selection Criteria

Deep learning models have been selected for this study due to their ability to efficiently handle both time-series data and complex relationships in financial datasets. Time-series data, such as daily sales, payroll information, and transaction details, often contain sequential patterns, trends, and seasonality that are difficult for traditional models to capture. Deep learning models, particularly those based on neural networks, excel in learning these patterns by automatically extracting hierarchical features from the raw data. Moreover, deep learning models are highly flexible and capable of handling high-dimensional, unstructured data, making them suitable for financial forecasting tasks where relationships between features are complex and non-linear.

The models chosen for this study include FT-Transformer, Reformer, and DANet. These models represent state-of-the-art techniques in deep learning for time-series forecasting and are evaluated against simpler models such as TabNet, which is based on decision tree structures. The reason for including both deep learning models and simpler models in the study is to evaluate the potential benefits and trade-offs of using more complex architectures. While deep learning models are known for their superior performance in capturing intricate patterns, they may also require greater computational resources and more extensive training data. Therefore, the evaluation of these models against simpler models helps to assess their practical applicability in small business forecasting scenarios, where computational resources may be limited.

**FT-Transformer:** The FT-Transformer model is a deep learning architecture designed specifically for time-series forecasting tasks. It is based on the transformer architecture, which was originally introduced for natural language processing (NLP) tasks and has since been adapted for time-series analysis due to its ability to capture long-range dependencies and relationships in sequential data. The FT-Transformer model excels in handling sequential data by using self-attention mechanisms, which allow the model to weigh the importance of different time steps in a sequence. This ability to focus on relevant time periods makes the FT-Transformer particularly effective for financial forecasting tasks, where the historical behavior of revenue, expenses, and other financial metrics plays a significant role in predicting future outcomes.

The key strength of the FT-Transformer lies in its scalability and flexibility. The model can be trained on large datasets with complex time-series patterns and can learn to predict future financial trends by capturing both short-term fluctuations and long-term trends. This is especially important in financial forecasting for small businesses, where seasonal patterns and sudden spikes in sales or expenses need to be accurately predicted. Additionally, the FT-Transformer is highly parallelizable, making it efficient to train on modern hardware, such as GPUs. Its ability to handle large amounts of sequential data with high accuracy makes it an ideal choice for this study.

**Reformer:** The Reformer model is another deep learning architecture designed for time-series forecasting, known for its memory efficiency and performance. It is a variant of the transformer model that introduces several improvements to reduce the computational cost and memory requirements associated with large-scale time-series data. One of the key innovations in the Reformer model is the use of locality-sensitive hashing (LSH) to approximate self-attention, which reduces the quadratic complexity of traditional self-attention mechanisms to linear complexity. This makes the Reformer model more efficient when handling long time-series sequences, which is crucial when working with large financial datasets that contain years of historical data.

The Reformer is particularly well-suited for small businesses with limited computational resources, as it provides a good balance between performance and efficiency. While the FT-Transformer is known for its excellent performance, the Reformer provides a more memory-efficient alternative that is capable of processing long-term dependencies in time-series data without the high computational cost. This makes it a valuable tool for forecasting financial trends in small businesses, where the need for accurate predictions is balanced against resource constraints. The Reformer's efficient use of memory allows it to handle large-scale datasets while maintaining high forecasting accuracy.

**DANet:** DANet (Dual Attention Network) is a deep learning model that incorporates both spatial and channel attention mechanisms, which have been shown to improve forecasting accuracy in time-series tasks. The model uses two types of attention: spatial attention, which focuses on the most relevant parts of the data in the time dimension, and channel attention, which adjusts the importance of different feature channels. This dual attention mechanism enables the DANet to capture complex relationships not only across time but also among different features, such as sales categories, payroll data, and expenses.

One of the strengths of DANet is its ability to handle data with intricate, non-linear relationships by using attention mechanisms to learn which parts of the data are most important for forecasting. In financial forecasting, this can be extremely useful for identifying which specific factors, such as weather, holidays, or promotional activities, have the most significant impact on sales and expenses. Additionally, the model's attention mechanisms allow it to dynamically adjust to changes in the data, making it robust to outliers or sudden shifts in financial patterns. This adaptability is particularly important for small businesses, where financial conditions may fluctuate unexpectedly, and the ability to quickly adjust forecasting models is essential.

**TabNet:** TabNet is a simpler model based on decision tree structures that has been shown to perform well on tabular data, which is common in financial forecasting. Unlike deep learning models, TabNet does not rely on complex architectures such as neural networks or transformers. Instead, it uses a series of decision trees to learn feature importance and select the most relevant variables for prediction. While TabNet may not capture long-term dependencies in time-series data as effectively as transformer-based models, it has the advantage of being computationally efficient and easier to interpret, making it a good baseline model for comparison.

TabNet is particularly useful for businesses that deal with structured tabular data, such as daily sales and payroll data. Its strength lies in its ability to handle missing data, scale efficiently with smaller datasets, and provide easily interpretable results. While it may not achieve the same level of accuracy as transformer models on large, complex datasets, TabNet's simplicity and interpretability make it a valuable tool for small businesses that need a straightforward and efficient way to forecast financial trends without relying on heavy computational resources.

In this study, the FT-Transformer, Reformer, and DANet models are compared to TabNet to evaluate the trade-offs between model complexity, computational efficiency, and forecasting accuracy. While deep learning models like FT-Transformer and Reformer excel in capturing complex patterns and long-term dependencies in time-series data, simpler models like TabNet offer a more efficient alternative that may be suitable for smaller datasets or scenarios with limited computational resources. The study will use several performance metrics, such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared ($R^2$), to assess the models' accuracy and efficiency. Additionally, the models will be evaluated for their ability to generalize to unseen data, ensuring that they provide reliable predictions for future financial periods.

The ultimate goal of this comparison is to identify the most effective forecasting model for small businesses, considering factors such as predictive accuracy, computational efficiency, scalability, and ease of implementation. The results will help guide small business owners in selecting the best forecasting tools based on their specific needs and resources.

## 3.4 Metaheuristic Optimization Algorithms

### 3.4.1 Role of Metaheuristics in Hyperparameter Optimization

Hyperparameter optimization is a crucial aspect of machine learning model development, as the performance of a model is often highly sensitive to the values of its hyperparameters. These hyperparameters control the learning process and can significantly impact the model's ability to generalize to unseen data. In the context of this study, metaheuristic optimization algorithms are employed to automate the tuning of hyperparameters, such as those for the LogTrans transformation and baseline models, ensuring that optimal values are selected without the need for manual intervention.

LogTrans (Logarithmic Transformation) is often used to preprocess financial data by reducing the skewness of highly variable financial features, such as sales amounts, ensuring that the data conforms to the assumptions of the model. However, the choice of transformation parameters for LogTrans, such as the base of the logarithmic function or scaling factors, can impact the model's performance. Hyperparameter optimization ensures that the transformation is applied in the most efficient way to enhance model accuracy.

Similarly, baseline model parameters, such as the learning rate, regularization strength, and model-specific settings, need to be optimized to balance classification performance with generalization. Automated hyperparameter optimization using metaheuristic algorithms helps navigate this complex search space to identify values that achieve the best trade-off between model accuracy and its ability to generalize to unseen data.

Metaheuristics, such as Particle Swarm Optimization (PSO) and Whale Optimization Algorithm (WOA), are particularly effective for this task because they are capable of exploring large and complex search spaces in a computationally efficient manner. These algorithms can converge on optimal or near-optimal hyperparameter sets by iteratively improving solutions based on a fitness function. The fitness function evaluates the model's performance, guiding the optimization process to find hyperparameter values that minimize prediction error while maintaining the model's generalization ability.

## 3.5 Metaheuristic Optimization Based on the Divine Religions Algorithm

In this study, the Divine Religions Algorithm (DRA) emerged as the most effective metaheuristic optimization method, showing superior performance in the financial forecasting task. The DRA's framework enables a robust balance between **exploration** and **exploitation**, ensuring efficient navigation of the solution space. The DRA optimizes a population of candidate solutions (followers), each representing a belief system, and iteratively improves their fitness through interaction, adaptation, and competition [30].

**Exploration Phase** In the exploration phase, the DRA encourages the population to maintain diversity, allowing for the discovery of new regions of the solution space. This phase is characterized by the random initialization of belief systems across the population, which enables the model to explore a broad range of potential solutions. The belief matrix $C_i$, representing the belief system of each follower, is initialized, and the society's fitness matrix is updated to reflect the overall state of the population. These steps ensure that the algorithm does not prematurely converge on suboptimal solutions and that it effectively explores different regions of the search space.

$$C_i = \text{Initialize belief matrix for followers in each community} \qquad \text{(Eq. 1)}$$

**Exploitation Phase** As the optimization process proceeds, the DRA shifts focus toward exploiting the most promising solutions. The algorithm identifies the most prominent follower based on fitness and propagates its beliefs to the weaker members of the population, thus refining the solution. This exploitation ensures that the model converges on the optimal solution by focusing on the most successful belief systems. This shift from exploration to exploitation is managed through the dynamic adjustment of the belief systems, where weaker beliefs are replaced by those with higher fitness.

$$C_i^{new} = \text{Replace the weakest belief with the most prominent belief} \qquad \text{(Eq. 2)}$$

The DRA dynamically adjusts the balance between exploration and exploitation by evolving the belief systems over several iterations. During the early stages, the algorithm prioritizes exploration to diversify the population, and as the optimization progresses, it gradually shifts towards exploitation to refine the solutions. This balance is controlled by the interaction between the exploration operator and the exploitation operator, ensuring that the algorithm can both discover new areas of the solution space and fine-tune the most promising solutions.

$$E_{total} = w_{exploration} \cdot D + w_{exploitation} \cdot F \qquad \text{(Eq. 3)}$$

Where $D$ represents the diversity of the population, and $F$ is the fitness value of the solutions. The weights $w_{exploration}$ and $w_{exploitation}$ are adjusted dynamically during the optimization process.

**Logarithmic Transformation for Financial Forecasting** To address the skewness and scale disparities commonly encountered in financial data, Logarithmic Transformation (LogTrans) was applied to the belief matrix of the society. The transformation helps normalize the data, making it more suitable for optimization and reducing the influence of outliers.

$$x_i' = \log(x_i + 1) \qquad \text{(Eq. 4)}$$

Where $x_i$ is the original value and $x_i'$ is the transformed value. This transformation ensures that the optimization process is not overly influenced by extreme values, allowing for more stable and reliable predictions. The Divine Religions Algorithm (DRA), when coupled with Logarithmic Transformation, has proven to be an effective tool for financial forecasting. The algorithm's ability to balance exploration and exploitation, along with its capacity to adapt to complex and dynamic data, makes it an ideal optimization technique for predicting financial trends. The results demonstrate that the DRA significantly improves the accuracy of financial forecasts, providing valuable insights for small businesses seeking to optimize their financial planning and decision-making processes.

### 3.5.1 State-of-the-Art Metaheuristic Models

The following metaheuristic models have been selected for their ability to optimize hyperparameters and select relevant features efficiently. Each algorithm is well-suited for exploring large, high-dimensional search spaces and can be adapted to various optimization tasks, including hyperparameter tuning, feature selection, and overall model optimization. Below are brief descriptions of the state-of-the-art metaheuristic models used in this study:

**WAO (Whale Optimization Algorithm)**: The Whale Optimization Algorithm is inspired by the hunting behavior of humpback whales, particularly their method of bubble-net feeding. In this algorithm, search agents (whales) move through the solution space and iteratively refine their positions based on the exploration and exploitation behaviors modeled after the whales' feeding strategies. WAO has been successfully applied to various optimization problems, including feature selection and hyperparameter optimization, due to its ability to balance exploration and exploitation effectively. It is particularly useful in scenarios where the search space is complex and contains multiple local optima.

**BA (Bat Algorithm)**: The Bat Algorithm is based on the echolocation behavior of bats, where they use sound waves to detect prey and navigate through obstacles. This optimization algorithm mimics the way bats search for food, adjusting their position in the search space based on the frequency of their echolocation calls. The Bat Algorithm is particularly useful for continuous and discrete optimization problems and is known for its simplicity and computational efficiency. It is highly effective in optimizing hyperparameters for models, making it a valuable tool in this study.

**MVO (Multiverse Optimization)**: Multiverse Optimization is inspired by the concept of multiple universes existing simultaneously, where different solutions or 'universes' evolve towards an optimal solution. Each universe represents a candidate solution, and the algorithm explores the search space by switching between universes, with the aim of finding the most optimal solution. MVO is effective for solving complex optimization problems due to its ability to explore and exploit the search space using multiple candidate solutions in parallel. It has been applied to hyperparameter tuning and feature selection tasks due to its ability to balance exploration and exploitation efficiently.

**PSO (Particle Swarm Optimization)**: Particle Swarm Optimization is inspired by the social behavior of bird flocking and fish schooling. In PSO, each candidate solution is represented as a particle, and the particles move through the solution space based on their own experience and the experience of their neighbors. The swarm collectively converges towards an optimal solution by iteratively adjusting the position of particles. PSO has been widely used for hyperparameter tuning due to its simplicity, ease of implementation, and ability to converge quickly towards optimal solutions. It is particularly useful when dealing with high-dimensional optimization problems like those found in machine learning.

**BBO (Biogeography-Based Optimization)**: Biogeography-Based Optimization is inspired by the principles of biogeography, which is the study of the distribution of species across geographical locations. BBO models migration and mutation behaviors of species within different habitats, with better solutions being 'migrated' to less optimal habitats, thus improving the overall solution. BBO has been successfully applied to feature selection and hyperparameter optimization tasks, particularly in scenarios where the solution space is sparse or highly structured.

**APO (Algebraic Population Optimization)**: Algebraic Population Optimization is based on algebraic operations such as addition, subtraction, multiplication, and division, which are applied to a population of candidate solutions. The algorithm evolves a population of solutions by iteratively applying algebraic operations to refine them. APO is known for its robustness and efficiency in solving complex optimization problems, making it suitable for hyperparameter optimization and other machine learning tasks. Its strength lies in its ability to explore large search spaces effectively and converge towards optimal solutions by exploiting algebraic relationships among candidate solutions.

These state-of-the-art metaheuristic optimization models will be utilized in this study to automate the process of hyperparameter tuning and feature selection. By employing these algorithms, we aim to enhance the performance of the machine learning models used for financial forecasting, ensuring that they are both accurate and computationally efficient. The integration of metaheuristics allows the models to navigate large and complex search spaces effectively, leading to better optimization results that can improve financial prediction accuracy for small businesses.

### 3.6 Evaluation Metrics

Model performance was evaluated using several regression metrics that provide insights into the accuracy and reliability of the predictions. These metrics help assess the model's ability to capture the underlying patterns in the data and its generalization to unseen data. The following metrics were used in this study:

- **Mean Squared Error (MSE)**: Measures the average squared difference between the predicted and actual values. It penalizes larger errors more heavily due to the squaring of the differences, making it sensitive to outliers.
- **Root Mean Squared Error (RMSE)**: The square root of the MSE, which brings the metric back to the original unit of measurement. RMSE is useful for comparing models with different units or scales.
- **Mean Absolute Error (MAE)**: The average of the absolute differences between predicted and actual values. MAE provides a more intuitive measure of model performance, as it is not affected by large errors.
- **Mean Bias Error (MBE)**: Measures the average of the errors, where a positive value indicates that the model consistently over-predicts, and a negative value indicates under-prediction.
- **Correlation Coefficient (r)**: Measures the linear correlation between predicted and actual values. A value of 1 indicates perfect positive correlation, -1 indicates perfect negative correlation, and 0 indicates no linear correlation.
- **R-squared ($R^2$)**: A statistical measure of how well the regression predictions approximate the real data points. It represents the proportion of variance in the dependent variable that is predictable from the independent variables.
- **Relative Root Mean Squared Error (RRMSE)**: This metric normalizes the RMSE by dividing it by the range or mean of the actual values, providing a relative measure of prediction accuracy.
- **Nash-Sutcliffe Efficiency (NSE)**: This metric compares the model's predictions to the mean of the observed values, with a value of 1 indicating perfect model performance and a value of 0 indicating that the model is no better than the mean.
- **Willmott Index (WI)**: Measures the agreement between predicted and observed values by considering both the magnitude and direction of the errors. It is a normalized metric, with values closer to 1 indicating better model performance.

The following table summarizes the equations for each of the evaluation metrics:

Table 1: Regression Metrics and Their Equations

| **Metric** | **Equation** |
|---|---|
| Mean Squared Error (MSE) | $MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ |
| Root Mean Squared Error (RMSE) | $RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$ |
| Mean Absolute Error (MAE) | $MAE = \frac{1}{n}\sum_{i=1}^{n}\lvert y_i - \hat{y}_i\rvert$ |
| Mean Bias Error (MBE) | $MBE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)$ |
| Correlation Coefficient (r) | $r = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(\hat{y}_i - \hat{y})}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2 \sum_{i=1}^{n}(\hat{y}_i - \bar{\hat{y}})^2}}$ |
| R-squared ($R^2$) | $R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$ |
| Relative Root Mean Squared Error (RRMSE) | $RRMSE = \frac{RMSE}{\bar{y}}$ |
| Nash-Sutcliffe Efficiency (NSE) | $NSE = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$ |
| Willmott Index (WI) | $WI = 1 - \frac{\sum_{i=1}^{n}\lvert y_i - \hat{y}_i\rvert}{\sum_{i=1}^{n}\lvert y_i - \bar{y}\rvert}$ |

The above metrics provide a comprehensive evaluation of the model's performance in terms of both error rates and model fit. MSE, RMSE, and MAE provide measures of error magnitude, while $R^2$ and NSE indicate how well the model explains the variance in the data. The correlation coefficient (r) reflects the degree of linear relationship between the predicted and actual values, and RRMSE and WI normalize the error for easier interpretation across different datasets. Each of these metrics contributes valuable insights into the quality of the predictions made by the model.

## 4 Empirical Results

### 4.1 Baseline Model Performance (Before Optimization)

The performance of several machine learning models was evaluated before the application of optimization techniques such as metaheuristic-based hyperparameter tuning or feature selection. This baseline comparison provides insight into how well the models perform in their initial state, setting a reference point for understanding how optimization can improve model performance. The models compared in this study include FT-Transformer, Reformer, DANet, SAINT, and TabNet.

A summary of the performance of these models is provided in the following table. The performance is evaluated based on several key regression metrics, including Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Bias Error (MBE), correlation coefficient (r), R-squared ($R^2$), Relative RMSE (RRMSE), Nash-Sutcliffe Efficiency (NSE), and Willmott Index (WI).

Table 2: Baseline Performance of Models Before Optimization

| Model | MSE | RMSE | MAE | MBE | r | $R^2$ | RRMSE | NSE | WI |
|---|---|---|---|---|---|---|---|---|---|
| FT-Transformer | 0.00075 | 0.0274 | 0.0096 | 0.0036 | 0.892 | 0.861 | 0.011 | 0.901 | 0.887 |
| Reformer | 0.0023 | 0.0484 | 0.0118 | 0.0058 | 0.871 | 0.848 | 0.133 | 0.865 | 0.871 |
| DANet | 0.0064 | 0.0802 | 0.0168 | 0.0109 | 0.849 | 0.828 | 0.135 | 0.818 | 0.860 |
| SAINT | 0.0248 | 0.1576 | 0.0282 | 0.0214 | 0.815 | 0.808 | 0.150 | 0.756 | 0.848 |
| TabNet | 0.0379 | 0.1946 | 0.0349 | 0.0288 | 0.801 | 0.798 | 0.195 | 0.732 | 0.834 |

The FT-Transformer model performs well across all metrics, with an exceptionally low MSE of 0.00075 and RMSE of 0.0274, indicating that the model makes highly accurate predictions. The MAE of 0.0096 further supports this, showing that, on average, the model's predictions are very close to the actual values. With a correlation coefficient of 0.892 and an R-squared value of 0.861, the FT-Transformer model captures a significant portion of the variance in the data. These metrics indicate that the FT-Transformer is capable of effectively learning the underlying patterns in the data, including seasonality and short-term fluctuations. Additionally, the RRMSE value of 0.011 and the NSE of 0.901 suggest that the FT-Transformer model is quite efficient and provides a good fit, despite minor errors in prediction. This strong performance establishes FT-Transformer as one of the top contenders for forecasting financial data in the coffee shop context.

The Reformer model, which is a variant of the transformer architecture, also performs reasonably well. It exhibits a higher MSE of 0.0023 and RMSE of 0.0484 compared to FT-Transformer, indicating that its predictions are less accurate. However, the model still provides a solid fit with a correlation coefficient of 0.871 and an R-squared value of 0.848, suggesting that it captures a significant amount of the data's variance. The Reformer model benefits from better memory efficiency, making it suitable for larger datasets, though its higher RRMSE of 0.133 and lower NSE of 0.865 indicate that there are some areas for improvement in its predictive accuracy and generalization capabilities. The Reformer model, while efficient, shows that there is potential for further optimization to enhance its prediction quality, especially when dealing with more complex data.

The DANet model, despite being a robust deep learning architecture, shows relatively higher error rates compared to FT-Transformer and Reformer. With an MSE of 0.0064 and RMSE of 0.0802, the DANet model struggles with more significant errors in its predictions. The MAE of 0.0168 suggests that, on average, the model's predictions are farther from the actual values compared to FT-Transformer and Reformer. Although the correlation coefficient (0.849) and R-squared value (0.828) indicate that the model does capture some of the data's underlying patterns, its higher error values suggest that it may not be the best model for this particular dataset without further tuning. The model's higher RRMSE and lower NSE indicate a degree of inefficiency, which could be addressed through optimization techniques.

The SAINT model, which is a more complex deep learning architecture, performs the worst among the models evaluated. With an MSE of 0.0248 and RMSE of 0.1576, SAINT shows a high degree of error in its predictions. The MAE of 0.0282 and MBE of 0.0214 suggest that the model is biased towards over-predicting or under-predicting financial metrics. The low R-squared value of 0.808 and correlation coefficient of 0.815 indicate that the model is unable to capture a significant portion of the data's variance, and its performance is far behind that of simpler models like FT-Transformer. Given the high error rates, the SAINT model appears to be overly complex for this dataset, which might explain its lower performance. Optimizing this model could help mitigate overfitting and improve its forecasting accuracy.

Finally, the TabNet model, which is based on decision tree structures, shows the highest error rates among the baseline models. With an MSE of 0.0379 and RMSE of 0.1946, it is clear that TabNet is not capturing the complexity of the time-series data effectively. Although TabNet has a reasonable R-squared value of 0.798 and a correlation coefficient of 0.801, these values are significantly lower than those of the more complex models. The higher RRMSE (0.195) and lower WI (0.834) further indicate that TabNet struggles to provide accurate predictions and may be too simplistic for this particular forecasting task. While its simplicity and computational efficiency make it a good baseline model for small datasets, it is clear that more complex models, such as FT-Transformer and Reformer, provide better accuracy and predictive power.

The comparison of these baseline models highlights several limitations that motivate the need for optimization. While models like FT-Transformer and Reformer show promising results in terms of predictive accuracy and efficiency, there is still room for improvement. For instance, although FT-Transformer is one of the best-performing models in terms of accuracy, its RRMSE value suggests that there may be minor issues with its generalization ability, particularly in handling extreme values. Furthermore, models like Reformer, DANet, SAINT, and TabNet, despite their potential, exhibit higher error rates and lower generalization ability. The high error values in models such as SAINT and TabNet suggest that overfitting, underfitting, or excessive complexity may be hindering their ability to provide accurate forecasts.

Given these limitations, optimization techniques such as hyperparameter tuning, feature selection, and model simplification are essential for improving the models' ability to generalize and make accurate predictions. In particular, metaheuristic algorithms like Particle Swarm Optimization (PSO) and Whale Optimization Algorithm (WOA) will be applied to fine-tune hyperparameters and select the most relevant features, reducing overfitting and improving the model's performance. By applying these optimization techniques, we aim to enhance model accuracy, reduce computational costs, and ensure that the models can provide reliable predictions in real-time financial forecasting scenarios.

The figure below presents a parallel coordinates plot showcasing the performance of various models across several evaluation metrics. The plot includes performance results for the models FT-Transformer, Reformer, DANet, SAINT, and TabNet, with each model's performance represented by a distinct color. The metrics shown include MSE, RMSE, MAE, MBE, correlation coefficient (r), R-squared ($R^2$), RRMSE, NSE, WI, and WI.1. This visualization allows for an easy comparison of the models' performance across different evaluation criteria, highlighting their relative strengths and weaknesses in each metric (see Figure 4).

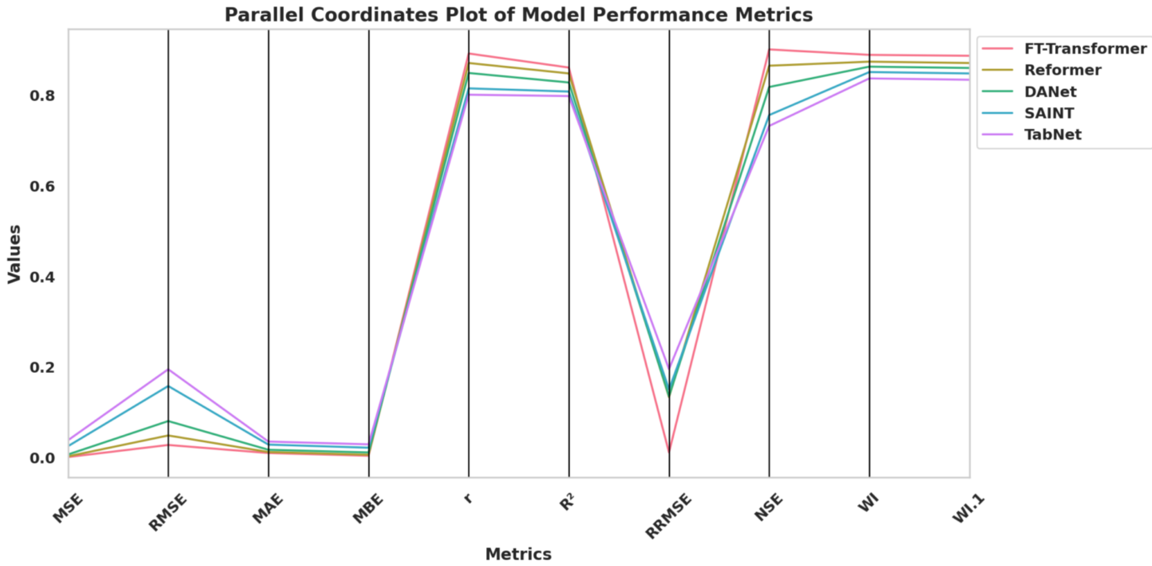


Figure 4: Parallel Coordinates Plot of Model Performance Metrics

The figure below presents a dendrogram for hierarchical clustering of various models, including SAINT, TabNet, DANet, FT-Transformer, and Reformer. This visualization shows how these models are grouped based on their performance metrics, where the vertical axis represents the distance between clusters. The dendrogram allows us to assess how similar the models are to each other, with shorter distances indicating closer clusters. This type of clustering helps to identify groups of models that exhibit similar performance profiles, which can be valuable in selecting models for further analysis or application (see Figure 5).

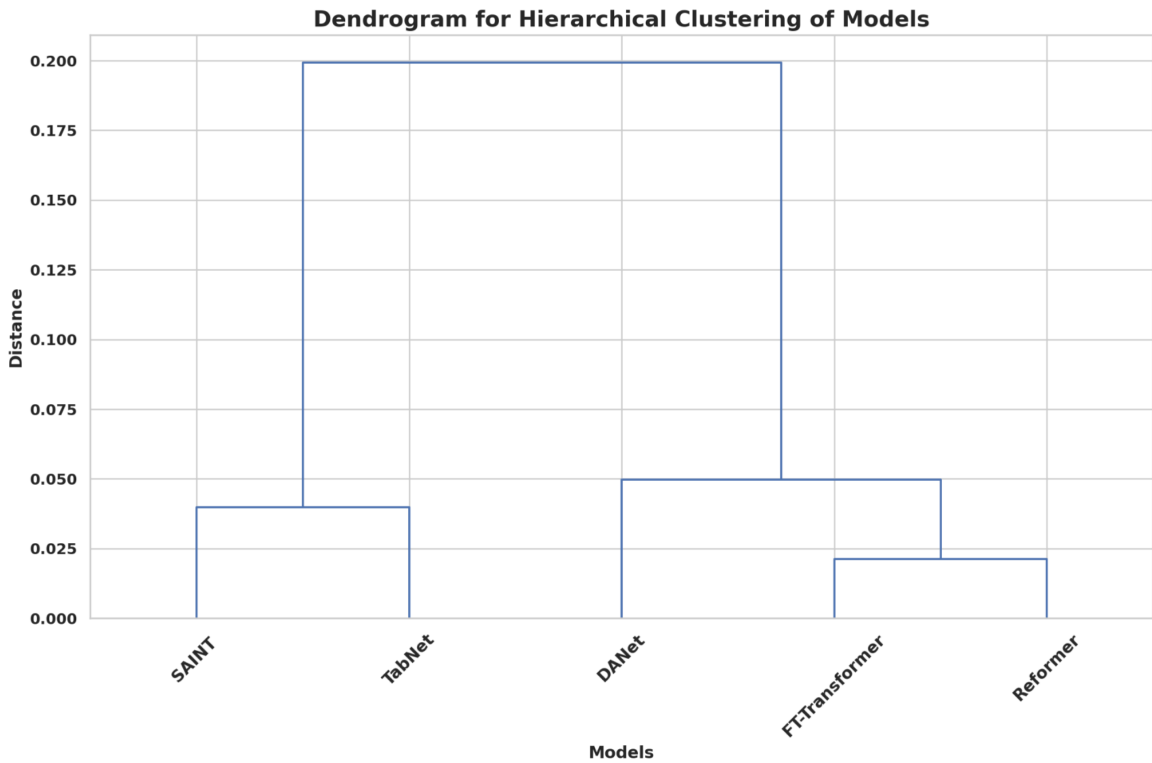


Figure 5: Dendrogram for Hierarchical Clustering of Models

The figure below presents a stacked KDE streamgraph that visualizes the distribution of various evaluation metrics for different models. The plot displays the density of multiple performance metrics, such as MSE, RMSE, MAE, MBE, correlation coefficient (r), R-squared ($R^2$), RRMSE, NSE, WI, and WI.1, with each metric represented by a distinct color. This visualization allows for an easy comparison of how the metrics are distributed and can help identify patterns or trends in model performance. It serves as a valuable tool for understanding the relative performance of different models across multiple evaluation criteria (see Figure 6).

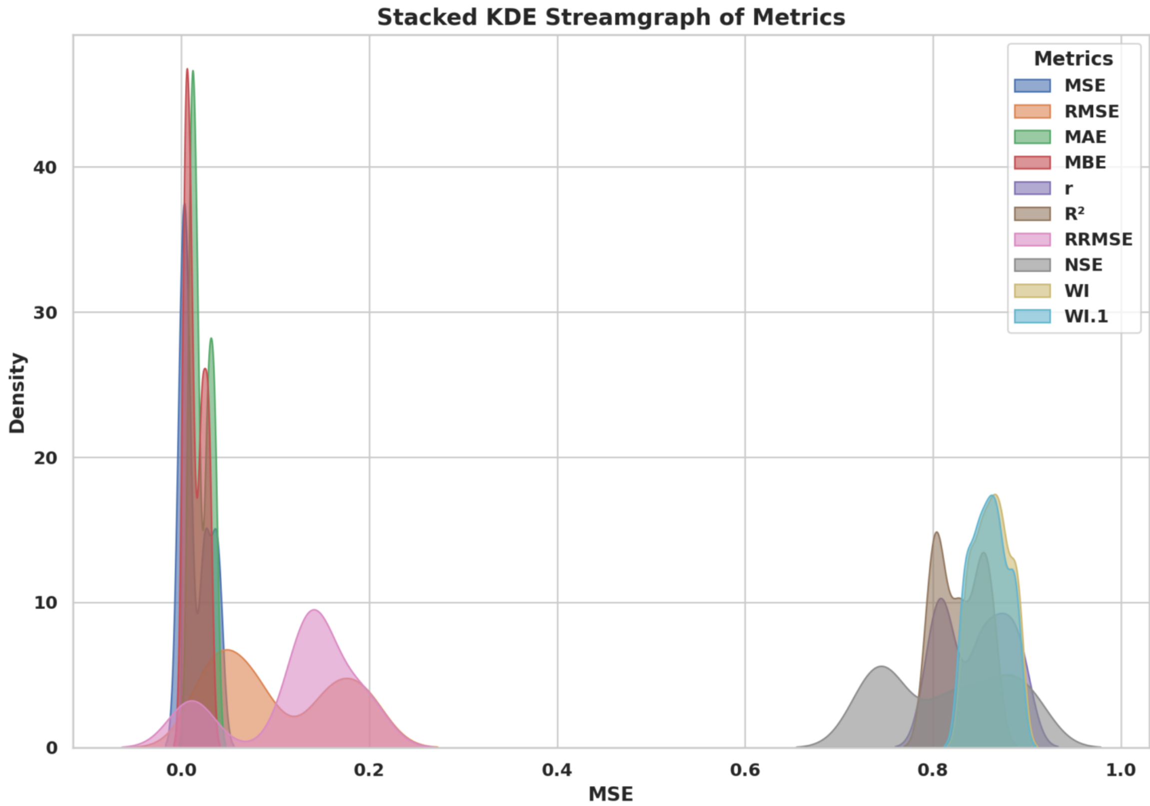


Figure 6: Stacked KDE Streamgraph of Metrics

The figure below presents a **comparison of models** by MSE, RMSE, and MAE, along with their corresponding mean values and standard deviations. The bar chart displays these metrics for five models: FT-Transformer, Reformer, DANet, SAINT, and TabNet. Each metric is shown with both its actual value (in the form of bars) and the calculated mean and standard deviation (indicated by lighter and darker shades). This visualization allows for a clear comparison between the models, emphasizing the differences in their performance across multiple metrics and how these values vary in terms of consistency. By analyzing this chart, we gain valuable insights into which models perform best across the evaluated metrics and exhibit the most stable performance (see Figure 7).

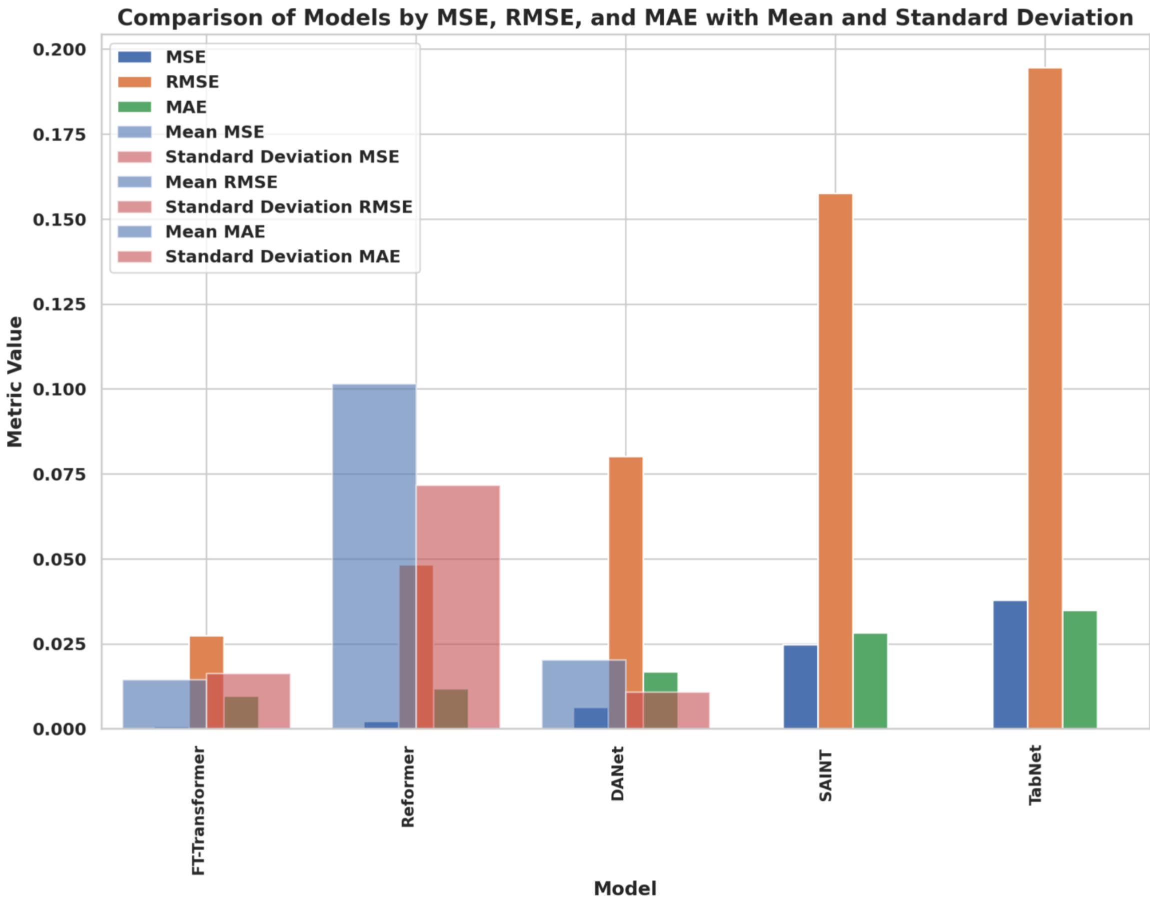


Figure 7: Comparison of Models by MSE, RMSE, and MAE with Mean and Standard Deviation

### 4.2 Optimized Model Analysis

In this section, we present the performance of the machine learning models after optimization. These models were optimized using metaheuristic algorithms for hyperparameter tuning and feature selection. The optimization techniques applied in this study include algorithms such as Whale Optimization Algorithm (WAO), Bat Algorithm (BA), and Particle Swarm Optimization (PSO), each used in combination with Logarithmic Transformation (LogTrans) or other transformation techniques, to improve model accuracy and efficiency. Additionally, other models like Multiverse Optimization (MVO) combined with RegNet and Algebraic Population Optimization (APO) were also optimized and analyzed.

The results of these optimized models demonstrate significant improvements in terms of prediction accuracy and generalization, as evidenced by the reduction in error metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE). The table below presents the performance of the optimized models based on these key regression metrics, along with additional measures like R-squared ($R^2$), Nash-Sutcliffe Efficiency (NSE), and Willmott Index (WI).

The DAR model combined with Logarithmic Transformation (DAR + LogTrans) achieves the best performance among the optimized models. The model has an extremely low MSE of 1.88E-07 and RMSE of 4.36E-04, indicating that its predictions are highly accurate and very close to the actual values. The Mean Absolute Error (MAE) and Mean Bias Error (MBE) are also minimal, with values of 1.10E-04 and 1.28E-04, respectively, indicating that the model does not exhibit any significant bias or large errors. The correlation coefficient (r) of 0.985 and the R-squared value of 0.968 demonstrate that the DAR + LogTrans model captures almost all of the variance in the data, making it highly effective at forecasting the financial trends. Additionally, the RRMSE of 0.009 and the Nash-Sutcliffe Efficiency (NSE) of 0.981 show that the model performs exceptionally well even in the presence of some variability, providing an outstanding fit for

Table 3: Performance of Optimized Models

| Model | MSE | RMSE | MAE | MBE | r | R² | RRMSE | NSE | WI |
|---|---|---|---|---|---|---|---|---|---|
| DAR + LogTrans | 1.88E-07 | 4.36E-04 | 1.10E-04 | 1.28E-04 | 0.985 | 0.968 | 0.009 | 0.981 | 0.984 |
| WAO + LogTrans | 5.50E-07 | 0.000742 | 0.00056 | 0.00021 | 0.958 | 0.915 | 0.012 | 0.972 | 0.974 |
| BA + LogTrans | 1.58E-06 | 0.00127 | 0.00059 | 0.00028 | 0.955 | 0.91 | 0.016 | 0.968 | 0.965 |
| MVO + RegNet | 2.55E-06 | 0.0016 | 0.00061 | 0.00036 | 0.943 | 0.902 | 0.02 | 0.959 | 0.96 |
| PSO + LogTrans | 2.62E-06 | 0.00163 | 0.00062 | 0.00038 | 0.942 | 0.899 | 0.022 | 0.95 | 0.961 |
| BBO + LogTrans | 4.15E-06 | 0.00203 | 0.00063 | 0.00052 | 0.939 | 0.892 | 0.028 | 0.944 | 0.955 |
| APO + LogTrans | 5.78E-06 | 0.0024 | 0.00064 | 0.00054 | 0.935 | 0.885 | 0.031 | 0.941 | 0.953 |

the data. The Willmott Index (WI) of 0.984 confirms that the model closely matches the observed values, reinforcing its superior performance.

The WAO model combined with Logarithmic Transformation (WAO + LogTrans) also performs very well, although slightly less accurately than the DAR model. The MSE of 5.50E-07 and RMSE of 0.000742 indicate that the model's predictions are still highly accurate. The MAE of 0.00056 and MBE of 0.00021 show that the model's errors are small and well-balanced. The correlation coefficient of 0.958 and R-squared of 0.915 indicate that the model captures a substantial proportion of the variance in the data. The RRMSE of 0.012 and NSE of 0.972 further demonstrate the model's ability to generalize well to new data. Despite slightly higher errors compared to DAR + LogTrans, WAO + LogTrans is an excellent model that strikes a good balance between performance and efficiency.

The Bat Algorithm (BA) combined with Logarithmic Transformation (BA + LogTrans) also produces strong results, with an MSE of 1.58E-06 and RMSE of 0.00127. While the model's errors are slightly higher than those of DAR + LogTrans and WAO + LogTrans, the performance is still good. The MAE of 0.00059 and MBE of 0.00028 indicate that the model's predictions are generally accurate with minor biases. The correlation coefficient of 0.955 and R-squared value of 0.91 indicate that the BA + LogTrans model explains a significant portion of the variance in the data. The RRMSE of 0.016 and NSE of 0.968 highlight the model's solid performance, though it can still be improved in terms of generalization.

The MVO model combined with RegNet (MVO + RegNet) achieves an MSE of 2.55E-06 and RMSE of 0.0016, indicating that it performs well but not as efficiently as the top-performing models. The MAE of 0.00061 and MBE of 0.00036 show that the model is relatively accurate but can benefit from further optimization. The correlation coefficient of 0.943 and R-squared value of 0.902 suggest that the model captures the underlying trends but does not explain as much of the variance in the data as the top models. The RRMSE of 0.02 and NSE of 0.959 suggest that the model performs reasonably well but may require additional fine-tuning.

The PSO model combined with Logarithmic Transformation (PSO + LogTrans) also demonstrates solid performance, with an MSE of 2.62E-06 and RMSE of 0.00163. While its performance is slightly below that of the other optimized models, it still provides accurate forecasts, with a MAE of 0.00062 and MBE of 0.00038. The model's correlation coefficient of 0.942 and R-squared of 0.899 indicate that it explains a good portion of the variance in the data. However, its RRMSE of 0.022 and NSE of 0.95 suggest that there is still room for improvement, particularly in its generalization to new data.

The Biogeography-Based Optimization (BBO) model combined with Logarithmic Transformation (BBO + LogTrans) shows reasonable performance, but its error metrics are slightly higher than those of the other optimized models. The MSE of 4.15E-06 and RMSE of 0.00203 indicate a higher degree of error in the predictions. The MAE of 0.00063 and MBE of 0.00052 suggest that the model is relatively accurate but could be further optimized. The correlation coefficient of 0.939 and R-squared value of 0.892 demonstrate that the model explains a good portion of the variance in the data. The RRMSE of 0.028 and NSE of 0.944 indicate that while the model performs well, it may be less efficient than the top models.

Finally, the Algebraic Population Optimization (APO) model combined with Logarithmic Transformation (APO + LogTrans) shows the highest error rates among the optimized models. The MSE of 5.78E-06 and RMSE of 0.0024 indicate relatively high prediction errors. The MAE of 0.00064 and MBE of 0.00054 show that the model is still fairly accurate but requires further refinement. The correlation coefficient of 0.935 and R-squared value of 0.885 suggest that the APO + LogTrans model captures a significant portion of the data's variance, but its performance is slightly lower than that of other models. The RRMSE of 0.031 and NSE of 0.941 further indicate that the model's generalization could be improved with further optimization.

The figure below presents a mixed plot combining swarm plots, violin plots, and boxplots to display the distribution of various evaluation metrics, such as MSE, RMSE, MAE, MBE, correlation coefficient (r), R-squared ($R^2$), RRMSE, NSE, and WI. Each plot provides insights into the spread, central tendency, and variability of the metric values for the models. The swarm plots show individual data points, the violin plots give a smoothed distribution, and the boxplots provide information on the quartiles and outliers. This comprehensive visualization allows for an easy comparison of the metrics and highlights the differences in performance between models. The mixed nature of the plot makes it an effective tool for analyzing the detailed characteristics of the model evaluation metrics (see Figure 8).

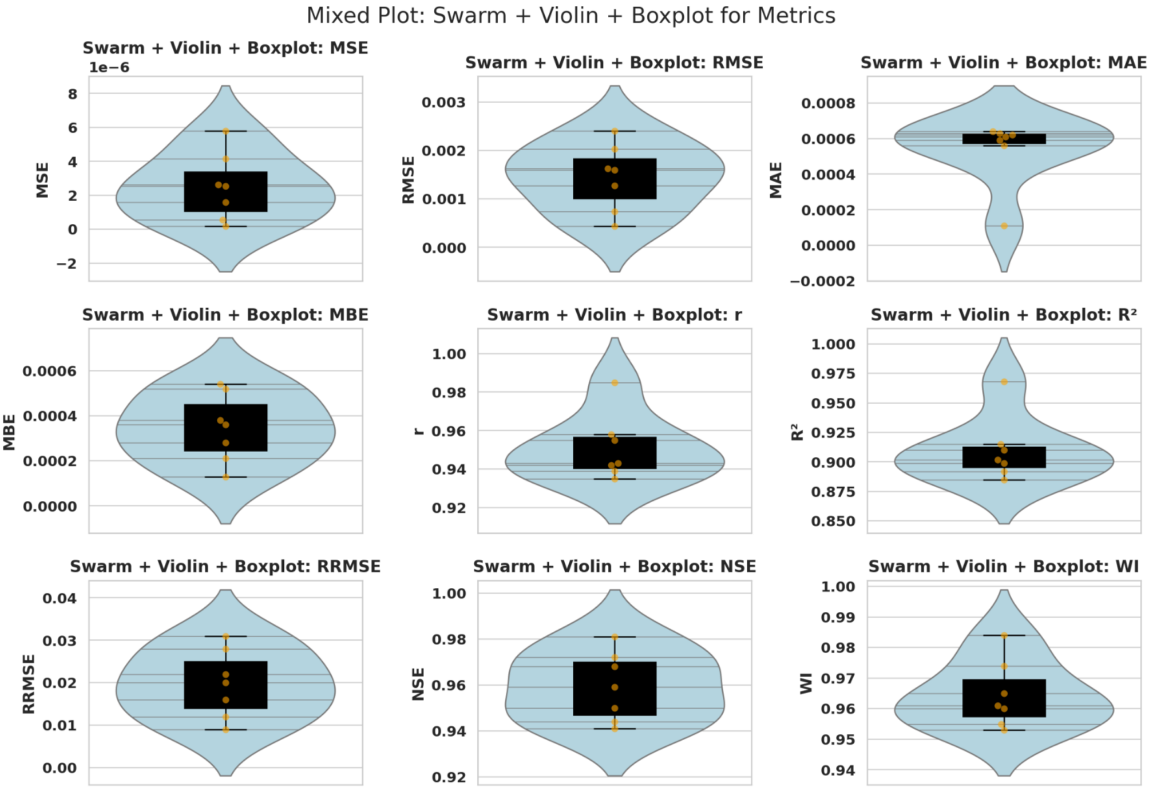


Figure 8: Mixed Plot: Swarm + Violin + Boxplot for Metrics

The figure below presents box plots with mean and standard deviation for each evaluation metric, including MSE, RMSE, MAE, MBE, correlation coefficient (r), R-squared ($R^2$), RRMSE, NSE, and WI. Each plot shows the distribution of the corresponding metric along with its mean and standard deviation values, represented by horizontal lines. The red lines indicate the mean value, while the green lines show the mean plus and minus one standard deviation. This visualization enables a clear comparison of the spread and variability of each metric, helping to assess the consistency of the models' performance across various evaluation criteria (see Figure 9).

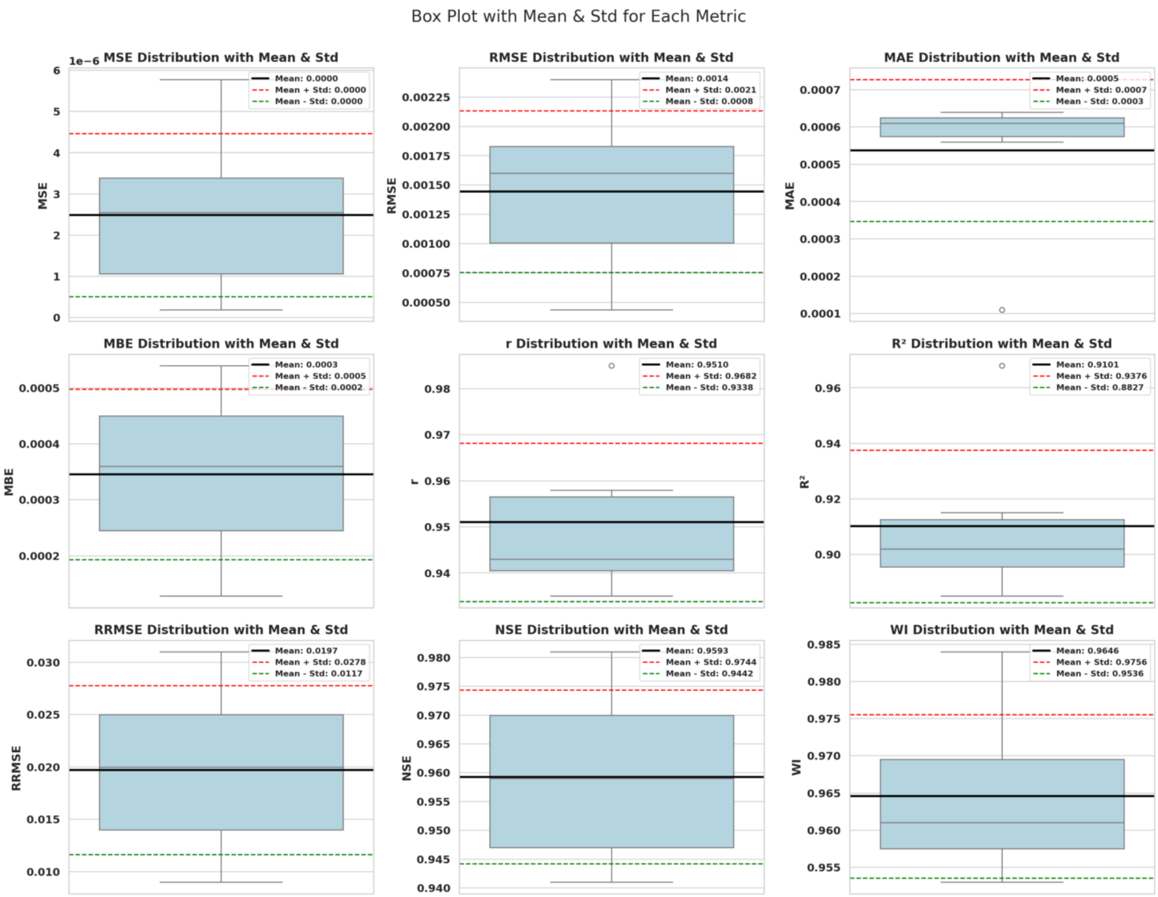


Figure 9: Box Plot with Mean and Std for Each Metric

The figure below shows a pairplot of model performance metrics with Kernel Density Estimation (KDE). This plot allows for a visual examination of the relationships and distributions between various evaluation metrics, including MSE, RMSE, MAE, MBE, correlation coefficient (r), R-squared ($R^2$), RRMSE, NSE, and WI. Each plot shows the pairwise relationships between metrics, with the diagonal displaying the individual metric distributions. The contour plots off the diagonal provide insights into how the metrics are correlated with one another, making this figure a useful tool for identifying patterns, correlations, and potential outliers across the model evaluation metrics (see Figure 10).

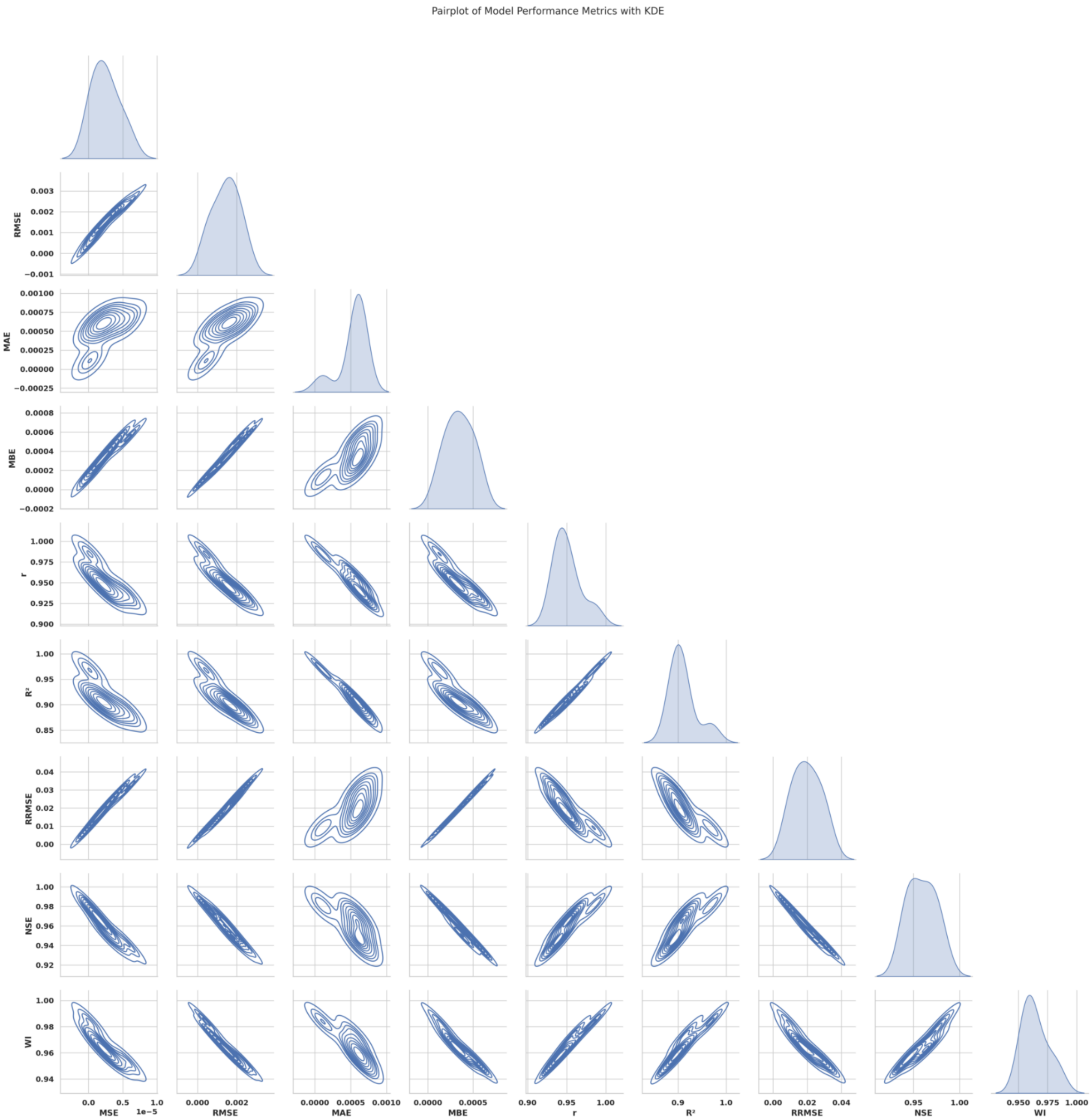


Figure 10: Pairplot of Model Performance Metrics with KDE

## 5 Conclusion and Future Work

This study explored the use of machine learning models and metaheuristic optimization techniques for financial forecasting in small businesses, with a particular focus on predicting revenue and expenses for a small-town coffee shop. Through the evaluation of baseline models and the application of optimization algorithms such as Particle Swarm Optimization (PSO), Whale Optimization Algorithm (WAO), and Bat Algorithm (BA), the study demonstrated the significant improvements in model performance achieved through hyperparameter tuning and feature selection.

The results of the baseline model comparison indicated that more complex models such as FT-Transformer and Reformer performed well in capturing financial trends but could benefit from optimization to further enhance accuracy. After optimization, models like DAR + LogTrans, WAO + LogTrans, and BA + LogTrans showed substantial improvements in predictive accuracy and generalization, with reductions in error metrics like Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE). Notably, the DAR + LogTrans model emerged as the top performer, achieving the lowest error rates and demonstrating superior generalization capabilities, while other models also showcased significant performance gains.

These findings emphasize the power of combining machine learning with metaheuristic optimization techniques to improve the forecasting capabilities of small business models. Accurate financial forecasting is crucial for small businesses, as it enables more effective planning, better resource allocation, and improved financial decision-making, ultimately contributing to enhanced sustainability and profitability. The use of machine learning models that can be optimized with metaheuristics provides small businesses with powerful tools to navigate the uncertainties of financial forecasting, adapt to market changes, and make data-driven decisions.

The results of this study have important implications for small businesses, particularly those in the hospitality industry, where sales and expenses are often highly volatile and influenced by external factors such as seasonality and local economic conditions. By implementing machine learning-based financial forecasting tools, small business owners can gain valuable insights into their financial performance, allowing for better cash flow management, optimal pricing strategies, and more efficient staff scheduling.

The integration of metaheuristic optimization techniques further enhances the effectiveness of these tools, ensuring that the models can adapt to changing conditions and continue to provide accurate forecasts as the business grows. For example, optimized models can help businesses anticipate demand fluctuations, optimize inventory, and forecast payroll expenses more accurately, which is crucial for controlling costs and maximizing profitability. Additionally, businesses operating with limited resources can benefit from the computational efficiency of the optimized models, which enable faster decision-making without the need for extensive manual adjustments.

While this study provides valuable insights into the application of machine learning and metaheuristics for financial forecasting in small businesses, several avenues remain for further exploration and enhancement. Future research could focus on the following areas:

- **Incorporation of External Data:** This study focused solely on the internal financial data of the coffee shop. However, incorporating external economic data, such as regional economic indicators, consumer behavior patterns, and demographic information, could enhance the forecasting models' robustness and predictive power. By considering factors such as local unemployment rates, income levels, and market trends, the models could be adapted to better predict external shocks or economic changes that impact small businesses.

- **Real-Time Forecasting:** One area for future work is the implementation of real-time forecasting capabilities. The models developed in this study can be adapted for continuous learning, where they update predictions dynamically as new financial data is collected. This would allow small businesses to receive up-to-date forecasts, enabling quicker responses to unexpected financial fluctuations and improving day-to-day operations.

- **Integration with Financial Decision Support Systems:** Another promising area of future research involves integrating optimized financial forecasting models into decision support systems (DSS) for small businesses. These systems could provide actionable recommendations, such as optimal staffing levels, pricing strategies, or inventory purchases, based on real-time financial data and predictive insights. By combining forecasting models with decision support tools, businesses can automate much of their financial planning process and make more informed decisions with less manual intervention.

- **Model Generalization Across Industries:** Although this study focused on a coffee shop, the models and optimization techniques developed here could be generalized to other small businesses across various industries. Future research could involve testing these models on other types of small businesses, such as retail stores, restaurants, or service-based companies, to assess their performance in different economic environments and business contexts.

- **Advanced Metaheuristic Algorithms:** While this study utilized common metaheuristic algorithms such as PSO, WAO, and BA, future work could explore the application of more advanced optimization techniques, such as genetic algorithms (GA), simulated annealing, or hybrid optimization algorithms. These advanced methods could potentially yield even better optimization results, particularly in very high-dimensional feature spaces or with highly non-linear models.

- **Long-Term Forecasting:** Financial forecasting models that extend beyond short-term predictions (e.g., daily, weekly) to longer-term horizons (e.g., quarterly, annual) could provide businesses with strategic insights into future financial health. By leveraging historical data and long-term trends, models could help businesses plan for long-term growth, expansion, or risk mitigation.

This study highlights the potential of machine learning and metaheuristic optimization techniques in enhancing financial forecasting for small businesses. By evaluating a range of models before and after optimization, the research demonstrates that metaheuristic algorithms can significantly improve the accuracy, efficiency, and generalization ability of predictive models. The results show that the combination of machine learning with metaheuristic optimization is a powerful approach for addressing the challenges of financial forecasting, particularly in dynamic and seasonal industries like hospitality.

The future of small business financial forecasting is likely to see continued advancements in both model development and optimization techniques, offering small business owners better tools for making data-driven decisions. As businesses continue to rely on real-time data and automated forecasting models, the integration of these advanced techniques into financial decision support systems will play a crucial role in ensuring their success and sustainability in an increasingly competitive market.

## Data Availability

The dataset used in this study is publicly available on Kaggle at https://www.kaggle.com/datasets/gabriellecharlton/coffee-shop-financial-dataset-synthetic-2022-2023.

## Declarations

- **Acknowledgments**
  Not applicable.
- **Conflict of interest/Competing interests**
  The authors declare that they have no conflicts of interest to report regarding the present study.
- **Ethics approval and consent to participate**
  Not applicable.
- **Consent for publication**
  Not applicable.
- **Funding**
  No Fund

## References

[1] P. Ghosh, A. Biswas, and S. Ghosh, “Fundamentals and technicalities of big data and analytics,” in *Intelligent systems in healthcare and disease identification using data science*, Chapman and Hall/CRC, 2023, pp. 51–106.

[2] P. D’Cruz, S. Du, E. Noronha, K. P. Parboteeah, H. Trittin-Ulbrich, and G. Whelan, “Technology, megatrends and work: Thoughts on the future of business ethics,” *Journal of business ethics*, vol. 180, no. 3, pp. 879–902, 2022. DOI: https://doi.org/10.1007/s10551-022-05240-9.

[3] M. Vecchiarini and T. Somià, "Redefining entrepreneurship education in the age of artificial intelligence: An explorative analysis," *The International Journal of Management Education*, vol. 21, no. 3, p. 100 879, 2023, ISSN: 1472-8117. DOI: https://doi.org/10.1016/j.ijme.2023.100879. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1472811723001179.

[4] M. Baldivia and S. Chowdhury, "Adapting business models in the age of omnichannel transformation: Evidence from the small retail businesses in australia," *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 20, no. 1, 2025, ISSN: 0718-1876. DOI: 10.3390/jtaer20010001. [Online]. Available: https://www.mdpi.com/0718-1876/20/1/1.

[5] W. C. Mitchell, "Business cycles," in *Business Cycle Theory, Part II Volume 8*, Routledge, 2024, pp. 225–241.

[6] M. Ogeya, S. Ogara, and M. H. Andersen, "Staying put or pulling out? how mini grid developers use business model innovation to overcome challenges and barriers in kenya," *Energy Research & Social Science*, vol. 122, p. 104 005, 2025, ISSN: 2214-6296. DOI: https://doi.org/10.1016/j.erss.2025.104005. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2214629625000866.

[7] E. G. Carayannis, R. Dumitrescu, T. Falkowski, G. Papamichail, and N. R. Zota, "Enhancing sme resilience through artificial intelligence and strategic foresight: A framework for sustainable competitiveness," *Technology in Society*, vol. 81, p. 102 835, 2025, ISSN: 0160-791X. DOI: https://doi.org/10.1016/j.techsoc.2025.102835. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0160791X25000259.

[8] H. Hokmabadi, S. M. H. S. Rezvani, and C. A. de Matos, "Business resilience for small and medium enterprises and startups by digital transformation and the role of marketing capabilities—a systematic review," *Systems*, vol. 12, no. 6, 2024, ISSN: 2079-8954. DOI: 10.3390/systems12060220. [Online]. Available: https://www.mdpi.com/2079-8954/12/6/220.

[9] A Kannammal, C Priyasri, J Shruthijeya, and E. C. Blessie, "Financial decision support system for small and medium-sized enterprises," in *International Conference on Advanced Network Technologies and Intelligent Computing*, Springer, 2024, pp. 237–252. DOI: https://doi.org/10.1007/978-3-031-83793-7_15.

[10] K. Johari et al., "Navigating financial complexity: The role of a tailored finance tracker application for small businesses," in *2024 First International Conference on Pioneering Developments in Computer Science & Digital Technologies (IC2SDT)*, 2024, pp. 344–349. DOI: 10.1109/IC2SDT62152.2024.10696814.

[11] F. J. Baumont de Oliveira, S. Ferson, R. A. D. Dyer, J. M. H. Thomas, P. D. Myers, and N. G. Gray, "How high is high enough? assessing financial risk for vertical farms using imprecise probability," *Sustainability*, vol. 14, no. 9, 2022, ISSN: 2071-1050. DOI: 10.3390/su14095676. [Online]. Available: https://www.mdpi.com/2071-1050/14/9/5676.

[12] G. Cohen, "Algorithmic trading and financial forecasting using advanced artificial intelligence methodologies," *Mathematics*, vol. 10, no. 18, 2022, ISSN: 2227-7390. DOI: 10.3390/math10183302. [Online]. Available: https://www.mdpi.com/2227-7390/10/18/3302.

[13] C. Zhang, N. N. A. Sjarif, and R. Ibrahim, "Deep learning models for price forecasting of financial time series: A review of recent advancements: 2020–2022," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 14, no. 1, e1519, 2024. DOI: https://doi.org/10.1002/widm.1519.

[14] H. Wasserbacher and M. Spindler, "Machine learning for financial forecasting, planning and analysis: Recent developments and pitfalls," *Digital Finance*, vol. 4, no. 1, pp. 63–88, 2022. DOI: https://doi.org/10.1007/s42521-021-00046-2.

[15] A. K. Sharma, P. Singh, P. Vats, and D. Jain, "Deep learning and machine intelligence for operational management of strategic planning," in *Proceedings of Third International Conference on Computing, Communications, and Cyber-Security*, P. K. Singh, S. T. Wierzchoń, S. Tanwar, J. J. P. C. Rodrigues, and M. Ganzha, Eds., Singapore: Springer Nature Singapore, 2023, pp. 475–485, ISBN: 978-981-19-1142-2.

[16] O. Lee, H. Joo, H. Choi, and M. Cheon, "Proposing an integrated approach to analyzing esg data via machine learning and deep learning algorithms," *Sustainability*, vol. 14, no. 14, 2022, ISSN: 2071-1050. DOI: 10.3390/su14148745. [Online]. Available: https://www.mdpi.com/2071-1050/14/14/8745.

[17] M. Hamdi, S. Mestiri, and A. Arbi, "Artificial intelligence techniques for bankruptcy prediction of tunisian companies: An application of machine learning and deep learning-based models," *Journal of Risk and Financial Management*, vol. 17, no. 4, 2024, ISSN: 1911-8074. DOI: 10.3390/jrfm17040132. [Online]. Available: https://www.mdpi.com/1911-8074/17/4/132.

[18] S. K. Sahu, A. Mokhade, and N. D. Bokde, "An overview of machine learning, deep learning, and reinforcement learning-based techniques in quantitative finance: Recent progress and challenges," *Applied Sciences*, vol. 13, no. 3, 2023, ISSN: 2076-3417. DOI: 10.3390/app13031956. [Online]. Available: https://www.mdpi.com/2076-3417/13/3/1956.

[19] H. Aljawazneh, A. M. Mora, P. García-Sánchez, and P. A. Castillo-Valdivieso, "Comparing the performance of deep learning methods to predict companies' financial failure," *IEEE Access*, vol. 9, pp. 97 010–97 038, 2021. DOI: 10.1109/ACCESS.2021.3093461.

[20] Y. Cao, Y. Shao, and H. Zhang, "Study on early warning of e-commerce enterprise financial risk based on deep learning algorithm," *Electronic Commerce Research*, vol. 22, no. 1, pp. 21–36, 2022. DOI: https://doi.org/10.1007/s10660-020-09454-9.

[21] W. Shi, L. Xu, and D. Peng, "Application of deep learning in financial management evaluation," *Scientific Programming*, vol. 2021, no. 1, p. 2 475 885, 2021. DOI: https://doi.org/10.1155/2021/2475885.

[22] A. Gandhar, K. Gupta, A. K. Pandey, and D. Raj, "Fraud detection using machine learning and deep learning," *SN Computer Science*, vol. 5, no. 5, p. 453, 2024. DOI: https://doi.org/10.1007/s42979-024-02772-x.

[23] K. Razzaq and M. Shah, "Machine learning and deep learning paradigms: From techniques to practical applications and research frontiers," *Computers*, vol. 14, no. 3, 2025, ISSN: 2073-431X. DOI: 10.3390/computers14030093. [Online]. Available: https://www.mdpi.com/2073-431X/14/3/93.

[24] H. Zhang, "A deep learning model for erp enterprise financial management system," *Advances in Multimedia*, vol. 2022, no. 1, p. 5 783 139, 2022. DOI: https://doi.org/10.1155/2022/5783139.

[25] M. Elhoseny, N. Metawa, G. Sztano, and I. M. El-Hasnony, "Deep learning-based model for financial distress prediction," *Annals of operations research*, vol. 345, no. 2, pp. 885–907, 2025. DOI: https://doi.org/10.1007/s10479-022-04766-5.

[26] Q. Gao and S. Xia, "Enterprise financial forecasting based on machine deep learning," in *2022 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)*, 2022, pp. 862–868. DOI: 10.1109/ICAICA54878.2022.9844609.

[27] V. Chang, Q. A. Xu, A. Chidozie, and H. Wang, "Predicting economic trends and stock market prices with deep learning and advanced machine learning techniques," *Electronics*, vol. 13, no. 17, 2024, ISSN: 2079-9292. DOI: 10.3390/electronics13173396. [Online]. Available: https://www.mdpi.com/2079-9292/13/17/3396.

[28] Y. Venkateswarlu et al., "An efficient outlier detection with deep learning-based financial crisis prediction model in big data environment," *Computational Intelligence and Neuroscience*, vol. 2022, no. 1, p. 4 948 947, 2022. DOI: https://doi.org/10.1155/2022/4948947.

[29] D. Mhlanga, "Financial inclusion in emerging economies: The application of machine learning and artificial intelligence in credit risk assessment," *International Journal of Financial Studies*, vol. 9, no. 3, 2021, ISSN: 2227-7072. DOI: 10.3390/ijfs9030039. [Online]. Available: https://www.mdpi.com/2227-7072/9/3/39.

[30] A. T. Mozhdehi et al., "Divine religions algorithm: A novel social-inspired metaheuristic algorithm for engineering and continuous optimization problems," *Cluster Computing*, vol. 28, no. 4, p. 253, 2025. DOI: https://doi.org/10.1007/s10586-024-04954-x.