Accurate Customer Financial Prediction Using Data-Driven Analytics in Retail Economics

Shahid Mahmood[1,*], Mahmoud Elshabrawy Mohamedr[2]

[1]School of Finance and Economics, Jiangsu University, Zhenjiang, People's Republic of China

[2]Computer Science and Intelligent Systems Research Center, Blacksburg 24060, Virginia, USA

Emails: shahidnajam786@live.com; mshabrawy@jcsis.org

Abstract

The growing availability of granular customer-level data has intensified the demand for accurate and robust predictive models in retail economics and consumer finance, particularly for forecasting financially relevant indicators such as savings capacity and credit-related measures, where prediction inaccuracies can lead to inefficient pricing strategies, misallocation of financial resources, and distorted risk assessments. Traditional statistical and econometric approaches often struggle to model the nonlinear and high-dimensional relationships inherent in such data, motivating the use of advanced deep learning techniques combined with intelligent optimization strategies. This study proposes an integrated economic and financial analytics framework that couples a sequence-to-sequence deep learning architecture (Sequence-to-Sequence, Seq2Seq) with state-of-the-art metaheuristic optimization algorithms for automated hyperparameter tuning, with particular emphasis on the Puma Optimizer–Seq2Seq (PO + Seq2Seq) configuration. The framework systematically evaluates multiple baseline deep learning models and enhances them through metaheuristic-driven optimization to address challenges related to convergence stability, generalization capability, and model sensitivity in customer-level financial prediction. Empirical analysis shows that the PO + Seq2Seq model consistently outperforms all baseline and alternative optimized configurations across all evaluation stages, achieving a Mean Squared Error of 2.05×10^{-5}, Root Mean Squared Error of 4.52×10^{-3}, Mean Absolute Error of 2.05×10^{-4}, and a very small Mean Bias Error of 5.40×10^{-5}, together with strong goodness-of-fit and efficiency indicators, including a correlation coefficient of 0.987, R^2 of 0.983, Nash–Sutcliffe Efficiency of 0.986, and Willmott Index of 0.988. From an economic and financial perspective, these findings demonstrate that the proposed PO + Seq2Seq framework provides a reliable and scalable predictive tool for customer analytics, enabling more accurate assessment of financial behavior, improved customer segmentation, and enhanced decision support in retail finance and consumer-oriented financial systems, while highlighting the critical role of metaheuristic optimization in unlocking the full predictive potential of deep learning models for real-world economic applications.

Keywords: Customer Financial Analytics; Retail Economics; Deep Learning Forecasting; Metaheuristic Optimization; Consumer Finance Prediction

1 Introduction

In recent decades, the rapid expansion in the availability and granularity of customer-level data has fundamentally transformed analytical practices in retail economics, consumer finance, and strategic marketing [1], [2], [3]. Firms operating in increasingly competitive and data-intensive environments can no longer rely on aggregate indicators or intuition-driven decision-making alone. Instead, economic value creation has become closely tied to the systematic exploitation of individual-level information that captures demographic characteristics, behavioral patterns, and financially relevant attributes [4], [5]. As a result, customer analytics has emerged as a cornerstone of modern economic and financial strategy, supporting more precise assessments of consumer heterogeneity, purchasing dynamics, and long-term relationship value [6], [7], [8].

From an economic standpoint, customer-level data provides essential microeconomic foundations for understanding consumption behavior, demand variability, and income–expenditure interactions [9], [10]. Such insights enable firms to better align pricing, inventory management, and promotional strategies with observed consumer behavior. In parallel, financial institutions increasingly rely on customer analytics to enhance credit assessment, savings estimation, and risk profiling processes. The ability to integrate behavioral indicators with financial characteristics allows for more refined evaluations of creditworthiness and financial resilience, which are critical for both profitability and financial stability.

Predictive modeling plays a central role within this analytical paradigm. Accurate prediction of customer-specific financial and behavioral indicators enables firms to anticipate future consumption trends, identify high-value customers, and mitigate potential financial risks [11], [12]. Forecasting variables such as spending behavior, estimated savings, and credit-related attributes contributes directly to revenue optimization, improved customer lifetime value estimation, and more efficient allocation of marketing and financial resources. In this context, prediction is not merely a technical task but a key economic mechanism that informs strategic planning and operational efficiency.

Machine learning has become an increasingly influential methodological framework for addressing these predictive challenges in financial analytics [13]. Traditional econometric and statistical methods, while grounded in economic theory, often impose restrictive assumptions regarding linearity, independence, and distributional form. These assumptions may be ill-suited for complex consumer datasets characterized by nonlinear relationships, heterogeneous effects, and intricate dependencies among demographic, behavioral, and financial variables. Machine learning models, by contrast, offer flexible, data-driven approaches capable of capturing complex patterns without requiring explicit functional specification.

The adoption of machine learning techniques has significantly enhanced the capacity of financial analytics systems to process high-dimensional data, extract latent behavioral signals, and adapt to evolving market conditions. By leveraging advanced learning architectures, machine learning models can uncover subtle interactions between income levels, spending habits, savings behavior, and customer loyalty that are difficult to identify using conventional approaches. Consequently, machine learning has demonstrated strong potential in improving predictive accuracy and decision support in consumer finance and retail economics.

Despite the methodological advantages offered by machine learning, the development of reliable customer-level financial prediction models remains subject to several important challenges. One of the most prominent issues arises from the high dimensionality of modern consumer datasets. The inclusion of diverse demographic attributes, behavioral indicators, and synthetic financial features increases the informational richness of the data but also introduces substantial computational and modeling complexity. High-dimensional feature spaces may lead to inefficient learning processes and increase the risk of overfitting, particularly when the available sample size is limited.

Feature redundancy and multicollinearity represent an additional challenge in financial customer analytics. Variables such as income, spending behavior, estimated savings, and credit-related measures are often intrinsically correlated due to underlying economic relationships. While these correlations reflect meaningful behavioral and financial dynamics, they can negatively affect model stability, interpretability, and generalization if not properly addressed. Redundant features may obscure the contribution of economically significant variables and inflate variance in predictive models, thereby undermining their reliability for financial decision-making.

Another critical challenge concerns the sensitivity of machine learning models to hyperparameter configurations. Hyperparameters control key aspects of the learning process, including model complexity, convergence behavior, and regularization strength. In the context of financial prediction, inappropriate hyperparameter choices can result in biased estimates, excessive variance, or unstable model behavior. Achieving an appropriate balance between bias and variance is particularly important in economic applications, where prediction errors can have material financial consequences.

Ensuring robust generalization performance further complicates the deployment of machine learning models in customer-level financial analytics. Models that perform well on historical data may fail to maintain accuracy when applied to new or evolving customer profiles. Such overfitting limits the practical usefulness of predictive systems in dynamic economic environments. Given the continuous evolution of consumer behavior and market conditions, developing models that generalize effectively to unseen data is essential for sustainable and reliable economic decision-making.

Motivated by the aforementioned challenges, this study aims to develop a comprehensive machine learning framework for customer-level financial prediction that emphasizes accuracy, robustness, and computational efficiency. The first objective is to systematically evaluate a set of advanced machine learning and deep learning models for predicting key financial indicators derived from enriched customer data. These indicators, including credit-related attributes and estimated savings, are of central importance in retail finance, customer segmentation, and strategic resource allocation.

A second objective is to enhance predictive performance through the application of metaheuristic optimization techniques. Metaheuristic algorithms provide flexible and efficient mechanisms for addressing complex optimization problems associated with feature selection and hyperparameter tuning. By integrating these algorithms into the modeling pipeline, the study seeks to identify compact and informative feature subsets, as well as optimal model configurations, that improve predictive accuracy while reducing computational overhead.

The final objective is to demonstrate that the proposed optimization-driven framework contributes to improved model robustness and generalization capability. By aligning methodological rigor with economic relevance, the study aims to provide predictive tools that support informed and reliable decision-making in consumer finance and retail economics.

This study contributes to the literature on economic and financial analytics in several important ways. First, it introduces an economically oriented feature selection framework designed to identify the most informative demographic, behavioral, and financial attributes influencing customer-level predictions. By emphasizing both predictive performance and economic interpretability, the proposed framework enhances the transparency and practical relevance of machine learning models.

Second, the study offers a structured comparative analysis of multiple machine learning approaches within a unified financial prediction context. Through consistent evaluation procedures and economically meaningful performance metrics, the analysis provides insights into the relative strengths and limitations of different modeling strategies for capturing complex consumer financial behavior.

Third, the study proposes an integrated optimization strategy that jointly addresses feature selection and hyperparameter tuning. This integrated perspective acknowledges the interdependence between data representation and model configuration, demonstrating how coordinated optimization can enhance predictive accuracy, computational efficiency, and interpretability. The resulting framework provides a scalable and adaptable solution for customer-level financial analytics in complex economic environments.

The remainder of this paper is organized as follows. The next section presents the dataset description, data preprocessing procedures, and methodological foundations of the machine learning and optimization approaches employed in this study. Subsequent sections provide a detailed empirical analysis and discuss the findings within an economic and financial context. The paper concludes with a summary of key insights, practical implications for consumer finance and retail economics, and directions for future research.

2 Literature Review

Decision-making, optimization, and data-driven intelligence have become central pillars of modern economic and financial systems, particularly as organizations operate in environments characterized by uncertainty, high dimensionality, and rapid technological change. The literature reviewed in this section spans operational research, decision theory, managerial economics, algorithmic governance, sustainability, and machine learning applications across logistics, finance, energy, agriculture, and digital markets. Collectively, these studies reflect a gradual but clear transition from static, deterministic, and assumption-heavy models toward adaptive, data-driven, and intelligent decision-support systems capable of handling complexity and uncertainty.

Early work in operations and logistics highlights the fragility of classical deterministic optimisation when confronted with stochastic demand. In the context of the dairy industry, the location-routing problem (LRP) is examined by comparing deterministic approaches based on sample-mean demand with stochastic ranking and selection methods, specifically indifference zone procedures [14]. The study demonstrates that using average demand estimates can systematically bias facility location and routing decisions, leading to suboptimal logistics performance. By contrast, indifference zone approaches explicitly account for uncertainty and prioritise probabilistic guarantees of correct selection while limiting computational cost. This contribution is significant in economic terms, as it shows how better modelling of uncertainty can directly improve operational efficiency, cost control, and service reliability in perishable goods distribution.

The conceptual foundations underlying such optimisation choices are deeply rooted in decision theory and decision science. A comprehensive exposition of the foundations of decision-making provides a philosophical and analytical framework for understanding rational choice, bounded rationality, and cognitive biases across economic and social systems [15]. By structuring decisions hierarchically—from routine operational choices to strategic and policy-level decisions—this body of work clarifies how decision quality depends on information availability, cognitive constraints, and contextual complexity. These insights are particularly relevant for economic and financial decision-making, where actors must evaluate trade-offs under uncertainty and incomplete information.

Decision-making complexity is further amplified in inter-organisational and relational contexts. Research on customer–supplier relationships in business markets emphasises that managerial decisions are rarely made in isolation but are embedded within interactive, high-involvement relationships [16]. This literature challenges the traditional axiomatic view of rationality, proposing instead an adaptive rationality framework in which decision-makers rely on heuristics shaped by experience, trust, and ongoing interaction. From an economic perspective, this shift explains how firms sustain long-term cooperative relationships despite uncertainty, incomplete contracts, and dynamic market conditions. The integration of applied psychology into industrial marketing research further enriches the understanding of decision-making capabilities as strategic economic assets.

As economic networks expand in scale and complexity, classical optimisation problems face new computational and structural challenges. The Uncapacitated Facility Location Problem (UFLP), a cornerstone of operations research, has been extended to network-based settings that explicitly incorporate demand intensity and topology [17]. The introduction of Network- and Demand-Weighted Roulette Wheel Initialization (NDWRWI) represents a methodological advance by embedding economic relevance—such as high demand and centrality—directly into the optimisation process. Empirical results show that intelligent initialisation significantly improves solution quality, clustering performance, and computational efficiency, particularly in ultra-large-scale networks. This work underscores the growing importance of hybrid heuristic–economic strategies in solving real-world optimisation problems.

Parallel to these methodological advances, the rise of algorithmic decision-making has transformed economic interactions, especially in data-intensive markets such as consumer credit. A multidisciplinary analysis of algorithmic decision-making in credit markets examines both economic efficiency and normative implications [18]. While increased data availability can improve risk assessment and contracting outcomes, the study highlights substantial risks related to bias, discrimination, model opacity, and strategic manipulation. Importantly, the literature connects economic costs with broader normative concerns, arguing that trust, privacy, and autonomy are essential components of sustainable algorithmic governance. This perspective is particularly relevant for financial regulation, where legal norms often evolve in response to economic theory and technological change.

Recent research increasingly integrates artificial intelligence into complex economic decision environments. The UNISONE framework exemplifies this evolution by combining generative artificial intelligence with multi-criteria decision reasoning and scenario-based simulation to support sustainable supply chain management [19]. Grounded in the Industry 5.0 paradigm, this model demonstrates how AI can enhance organisational agility, carbon efficiency, and sourcing resilience. From an economic standpoint, the framework illustrates how intelligent decision-support systems can reconcile profitability with long-term sustainability objectives, addressing a critical tension in modern supply chains.

The strategic embedding of analytics into decision processes is further formalised through the concept of data-driven decision-making (DDDM) [20]. This literature argues that the true economic value of data science lies not merely in prediction accuracy but in its integration across the full decision lifecycle, including preparation, execution, and evaluation. By distinguishing between programmed and non-programmed decisions, this work highlights how DDDM supports both operational efficiency and strategic innovation, linking analytics adoption to data entrepreneurship and competitive advantage.

Logistics and distribution systems provide a concrete setting in which AI-driven decision-making delivers measurable economic benefits. In last-mile delivery (LMD), deep neuroevolutionary algorithms have been proposed to decompose complex, integrated scheduling problems into tractable configuration-based models [21]. By coupling predictive deep learning with optimisation feedback, these approaches significantly outperform traditional prescriptive models in terms of computational efficiency, scalability, and solution quality. The economic implications are substantial, as last-mile delivery costs represent a significant share of total logistics expenditure in e-commerce and retail.

Behavioural dimensions of economic decision-making are further explored through studies of consumption under constraint. An investigation into counterfeit consumption among Bottom-of-the-Pyramid consumers reveals how bounded rationality, normative influence, and vulnerability interact to shape purchasing behaviour [22]. The identification of a "zone of awareness–indifference" provides a nuanced understanding of how economic decisions are made when consumers face limited information and constrained choice sets. These insights have implications for marketing strategy, consumer protection policy, and inclusive economic development.

Sustainability-oriented decision-making is also addressed through optimisation models for e-waste management. A multi-objective mixed-integer linear programming model integrated with goal programming demonstrates how economic, environmental, and operational objectives can be simultaneously optimised [23]. Compared with traditional heuristics, the proposed approach achieves substantial cost reductions while improving resource recovery and reducing environmental impact. This study reinforces the economic feasibility of circular economy strategies when supported by robust decision-support tools.

At the macroeconomic level, the relationship between digital economy development and energy productivity has been examined using artificial neural networks and explainable AI techniques [24]. The findings reveal non-linear, regionally heterogeneous effects, with digital inclusive finance emerging as a dominant driver over time. These results highlight the role of digitalisation as a catalyst for productivity growth and energy efficiency, offering important insights for economic and energy policy design in developing and emerging economies.

Machine learning applications further extend into agriculture and financial markets. Comparative studies in agricultural production forecasting demonstrate that traditional machine learning models, such as Random Forests, can outperform deep learning approaches when data are limited or structured [25]. Conversely, in financial markets, deep learning architectures—including ANN, CNN, LSTM, and hybrid CNN–LSTM models—consistently achieve superior predictive performance by capturing complex temporal dependencies and non-linear patterns [26], [27]. These contrasting results emphasise the importance of aligning model complexity with data characteristics in economic forecasting tasks.

Finally, financial risk management in e-commerce enterprises has emerged as a critical application domain for deep learning-based early warning systems [28]. By modelling financial risk as a dynamic and evolving process, these systems enable earlier detection of distress signals, supporting proactive intervention and reducing the likelihood of bankruptcy. From a financial economics perspective, such models enhance firm-level resilience and contribute to broader market stability.

In summary, the expanded body of literature demonstrates a clear convergence toward adaptive, intelligent, and data-driven decision-making across economic and financial domains. Whether addressing logistics optimisation, consumer behaviour, sustainability, or financial forecasting, these studies collectively underscore the necessity of integrating uncertainty modelling, behavioural insights, and artificial intelligence to support robust and effective economic decisions in increasingly complex environments.

3 Materials and Methods

3.1 Dataset Description

The empirical foundation of this study is an enhanced version of the widely used Mall Customers dataset, which has been systematically extended to overcome the limitations of its original formulation and to better reflect the characteristics of real-world economic and financial data. The original dataset was primarily designed for illustrative clustering exercises and contained only a small set of basic demographic variables, such as age and gender, alongside annual income and an abstract spending score. While adequate for introductory unsupervised learning demonstrations, this restricted feature space limited its applicability for deeper economic analysis and financially meaningful prediction tasks.

The enhanced dataset employed in this study introduces a richer and more realistic set of variables that are commonly encountered in retail economics, consumer finance, and marketing analytics. These additional features are synthetically generated but constructed to be logically consistent with observed income and spending patterns, thereby preserving economic plausibility. The motivation behind this enhancement is to bridge the gap between simplified educational datasets and the complex, multidimensional data environments faced by practitioners and researchers in applied financial analytics. As such, the dataset provides an appropriate testing ground for advanced machine learning methodologies while maintaining interpretability and economic relevance.

From an economic standpoint, the dataset is well suited for analyzing consumer heterogeneity and individual-level financial behavior. The integration of demographic, behavioral, and financial attributes enables the examination of income–consumption relationships, savings behavior, and customer engagement dynamics within a unified analytical framework. This structure supports a range of economically relevant tasks, including customer segmentation, financial capacity assessment, and predictive modeling of credit-related indicators. Consequently, the dataset aligns closely with real-world applications in retail finance, customer relationship management, and data-driven marketing strategy.

The feature set is organized into three broad categories. The first category consists of demographic features, including customer age, gender, and age group. Age is recorded as a continuous variable measured in years, while age group is derived through binning into discrete intervals to facilitate demographic segmentation and to capture nonlinear age-related effects. Gender is included as a categorical attribute, reflecting its role as a fundamental demographic characteristic in consumption and financial behavior analysis. Together, these variables provide essential contextual information that underpins customer-level economic modeling.

The second category encompasses economic and financial features that directly reflect the customer's financial position. Annual income, expressed in thousands of monetary units, serves as a primary indicator of purchasing power and economic capacity. Estimated savings is introduced as a derived variable, constructed from income and spending behavior to approximate the customer's ability to accumulate financial reserves. The credit score variable, although synthetic, is designed to emulate realistic credit assessment mechanisms by incorporating information related to income stability and consumption patterns. These financial attributes are central to the study's predictive objectives, as they capture key dimensions of financial resilience, risk exposure, and future consumption potential.

The third category includes behavioral features that characterize customer engagement and consumption preferences. The spending score represents a relative measure of spending intensity, capturing differences in consumption behavior across customers. Loyalty years approximates the duration of the customer's relationship with the retailer, providing a proxy for long-term engagement and relationship value. Preferred category is a categorical variable that reflects simulated shopping preferences across different consumption segments, such as luxury-oriented or budget-oriented behavior. These behavioral indicators are essential for linking financial capacity with observed consumption patterns and for supporting economically meaningful customer segmentation.

The predictive task considered in this study is formulated as a supervised learning problem focused on a financially relevant target variable. Depending on the specific analytical objective, the target variable is defined as either the customer's credit score or estimated savings. Both targets are of significant economic importance, as they summarize key aspects of financial stability and long-term value creation. Credit score prediction is closely aligned with risk assessment and credit allocation decisions, while estimated savings prediction supports analyses related to financial resilience, consumption smoothing, and customer lifetime value.

To ensure methodological rigor and robust evaluation, the dataset is partitioned into three non-overlapping subsets: training, validation, and testing sets. The training set is used for model estimation, allowing learning algorithms to capture underlying patterns in the data. The validation set supports model selection, feature selection, and hyperparameter tuning, ensuring that modeling decisions are not biased by test data exposure. The testing set is reserved exclusively for out-of-sample evaluation, providing an unbiased assessment of generalization performance. This three-way partitioning strategy is essential in economic and financial prediction contexts, where reliable out-of-sample performance is critical for real-world decision-making.

To establish an initial understanding of the market dynamics embedded in the dataset, we conduct an exploratory visual analysis of key price and volume variables. This step is essential for identifying fundamental statistical properties, inter-variable relationships, and potential irregularities that may influence subsequent modeling and forecasting tasks. In particular, the analysis focuses on the distributional characteristics of closing prices, the linear association between opening and closing prices, the comparative dispersion of open, high, low, and close (OHLC) values, and the density structure of trading volume. As illustrated in Figure 1, these visual diagnostics provide complementary perspectives on price behavior and market activity. The scatter-based representations highlight the strength and consistency of price movements, while the boxplots and density estimates reveal skewness, variability, and the presence of extreme observations. Collectively, these insights motivate appropriate data transformations and robust modeling strategies in later stages of the study.

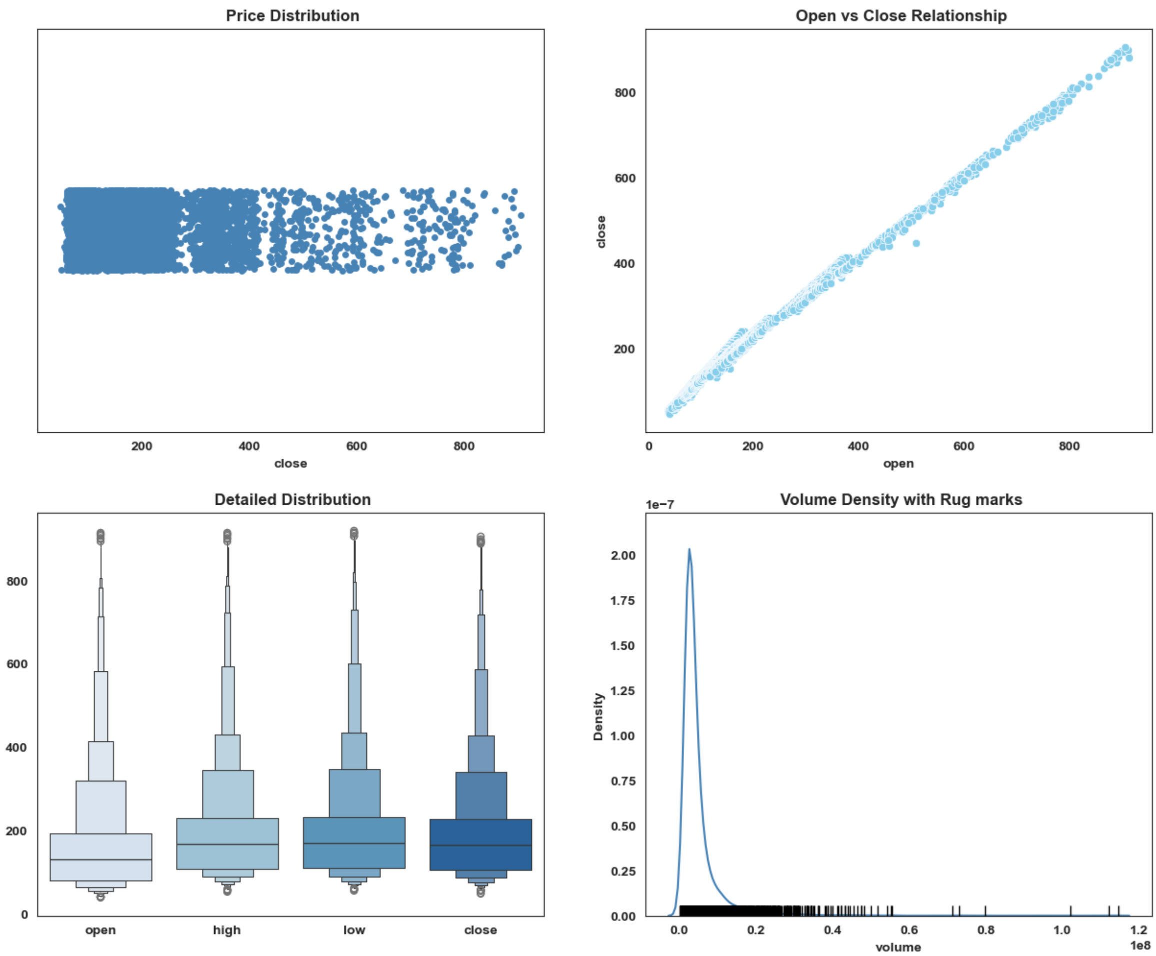

Figure 1: Exploratory visualization of financial time-series data, including price distributions, open–close relationships, OHLC boxplots, and volume density with rug marks.

To examine the long-term evolution and volatility characteristics of Goldman Sachs' equity performance, we analyze the annual distribution of closing prices over an extended historical horizon. This visualization is particularly useful for capturing structural changes in price behavior across different market regimes, including periods of growth, stagnation, and heightened uncertainty. As shown in Figure 2, the year-wise boxplots summarize the central tendency, dispersion, and presence of extreme values in closing prices, thereby enabling a comparative assessment of inter-annual variability. The progression of medians and the widening of interquartile ranges in later years reflect both sustained price appreciation and increasing volatility, while anomalous spreads in specific years highlight the impact of major financial events. Overall, this figure provides critical insights into the temporal dynamics of asset valuation and serves as a foundation for subsequent econometric modeling and risk analysis.

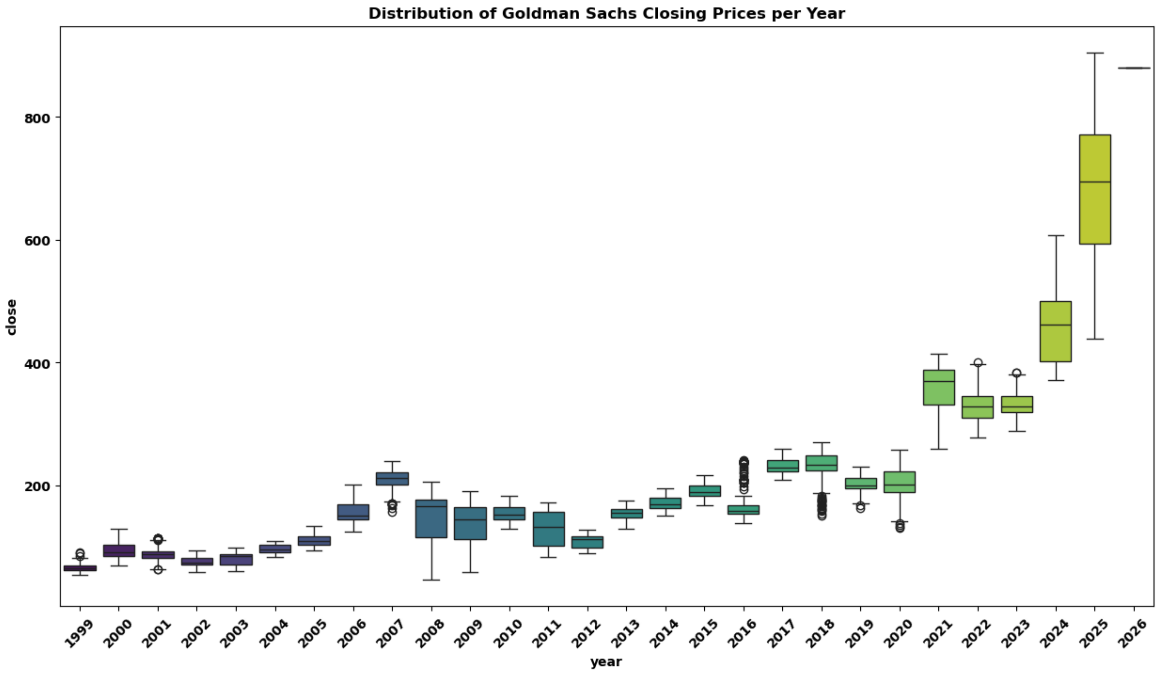

Figure 2: Distribution of Goldman Sachs closing prices by year, illustrating annual variations in central tendency, dispersion, and outliers over the observed period.

To further characterize the distributional properties of asset prices across both long-term and short-term horizons, we employ nonparametric density estimation and high-resolution scatter-based visualization. This analysis enables a nuanced examination of price concentration, multimodality, and dispersion patterns that are not readily observable through summary statistics alone. As illustrated in Figure 3, the kernel density estimate captures the overall shape of closing price distributions from 1999 to 2025, highlighting dominant price regimes and the gradual emergence of higher-price clusters over time. Complementarily, the swarm plot provides a granular view of daily closing prices in 2025, revealing the extent of intra-year variability and the density of observations around key price levels. Together, these visual representations offer insight into both historical price evolution and recent market behavior, thereby informing subsequent volatility modeling and predictive analysis.

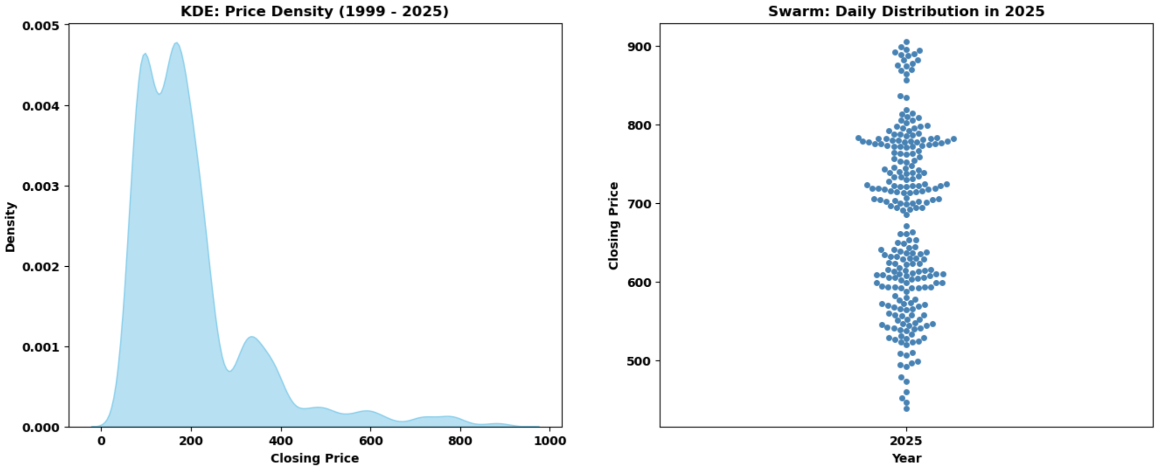

Figure 3: Kernel density estimation of closing prices over the full sample period and swarm plot illustrating the daily distribution of closing prices in 2025.

3.2 Data Preprocessing

A comprehensive data preprocessing pipeline is implemented prior to model development to ensure data integrity, numerical stability, and compatibility with machine learning algorithms. Given the inclusion of both observed and synthetically generated variables, particular attention is paid to verifying consistency across feature distributions and maintaining economically meaningful relationships among variables. The preprocessing stage plays a crucial role in mitigating noise, reducing bias, and enhancing the reliability of subsequent predictive modeling.

The first step of preprocessing involves the identification and treatment of missing values. Although the dataset is designed to be complete, missing or anomalous values may arise due to data generation processes or transformation steps. Appropriate imputation strategies are applied to address such cases while preserving the statistical properties of the data. The choice of imputation method is guided by the nature of each variable and its economic interpretation, ensuring that imputed values do not distort underlying income–spending or financial relationships.

Categorical features, including gender, age group, and preferred category, are transformed into numerical representations suitable for machine learning models. Encoding methods are selected to balance information preservation with computational efficiency, avoiding excessive dimensional expansion that could exacerbate the curse of dimensionality. The transformation of categorical variables is conducted in a manner that retains their economic meaning and supports interpretability in downstream analysis.

Feature scaling and normalization constitute another essential preprocessing step, particularly given the heterogeneous scales of economic and financial variables. Annual income, estimated savings, and credit-related measures may differ substantially in magnitude, which can negatively affect learning dynamics and optimization stability if left unaddressed. Standardization and normalization techniques are therefore applied to align feature distributions, facilitate efficient convergence during model training, and ensure that no single variable disproportionately influences the learning process due to scale effects.

Finally, a detailed correlation and redundancy analysis is performed to examine interdependencies among features. Economic variables such as income, spending score, estimated savings, and credit score are inherently related, reflecting underlying behavioral and financial mechanisms. While these relationships are economically meaningful, excessive multicollinearity can impair model stability and reduce interpretability. Correlation analysis is used to identify highly correlated feature pairs, and redundancy assessments guide subsequent feature selection efforts. By systematically addressing correlation and redundancy at the preprocessing stage, the study establishes a robust foundation for effective feature selection and reliable financial prediction in later stages of the analysis.

3.3 Deep Learning Models

The selection of deep learning models in this study is guided by their demonstrated capability to handle complex nonlinear relationships, heterogeneous feature spaces, and sequential dependencies that commonly arise in economic and financial data. Customer-level financial datasets often exhibit intricate interactions between demographic characteristics, behavioral indicators, and financial attributes, as well as latent temporal or sequential structures related to customer engagement and consumption dynamics. Deep learning architectures are particularly well suited to such settings, as they provide flexible function approximation capabilities and can adapt to high-dimensional input spaces without requiring strong parametric assumptions.

The model selection process emphasizes robustness, scalability, and suitability for financial prediction tasks rather than purely algorithmic novelty. Priority is given to models that have been widely adopted in forecasting and sequence modeling contexts and that are capable of learning complex patterns from structured data. By employing a diverse set of deep learning architectures, the study aims to capture complementary modeling strengths and to establish a comprehensive benchmarking framework for customer-level financial prediction.

The baseline modeling framework includes several advanced deep learning architectures, namely Seq2Seq, N-HITS, Informer, NODE, and GRU. These models are selected to represent different design philosophies and learning mechanisms, ranging from sequence-to-sequence learning and hierarchical forecasting structures to attention-based architectures and neural differential modeling. The inclusion of multiple architectures allows for a systematic examination of how different learning paradigms perform when applied to enriched customer data with economic and financial relevance.

Seq2Seq models are included due to their strong capacity for mapping input sequences to output sequences, making them particularly suitable for structured prediction tasks where complex dependencies exist among input features. In the context of customer analytics, Seq2Seq architectures can effectively model interactions between financial indicators and behavioral attributes, enabling flexible representation of nonlinear relationships that influence customer-level financial outcomes.

N-HITS is incorporated as a representative of hierarchical deep learning approaches designed to capture multiscale patterns in structured data. Such architectures are well suited for financial and economic datasets where patterns may emerge at different levels of aggregation or abstraction. By decomposing complex signals into hierarchical components, N-HITS can enhance learning efficiency and improve the representation of underlying financial dynamics.

Informer models are included to account for attention-based learning mechanisms that have proven effective in handling long-range dependencies and large input spaces. In financial prediction tasks, attention mechanisms allow models to selectively focus on the most informative features or temporal segments, which is particularly valuable when dealing with high-dimensional customer data. This selective focus supports improved learning efficiency and interpretability in complex economic settings.

NODE models represent a distinct modeling paradigm that treats deep learning as a continuous transformation process. Their inclusion reflects the growing interest in neural differential approaches for modeling smooth and complex relationships in financial data. Such models are particularly relevant for capturing gradual changes in customer financial behavior and for providing flexible representations that align with continuous economic processes.

GRU models are incorporated due to their established effectiveness in sequence modeling and time-dependent prediction tasks. Their gated structure enables efficient learning of dependencies while mitigating issues related to vanishing gradients. In customer-level financial analytics, GRU architectures are well suited for capturing sequential patterns related to loyalty, spending behavior, and evolving financial characteristics, even when explicit time-series structure is limited or implicit.

Overall, the selected deep learning models collectively provide a robust and diverse modeling foundation for financial prediction tasks. Their relevance to customer analytics lies in their ability to learn complex nonlinear relationships, handle heterogeneous feature sets, and adapt to structured and sequential data commonly encountered in economic and financial applications. By establishing these models as baselines, the study creates a solid reference point for subsequent optimization and enhancement through advanced feature selection and hyperparameter tuning strategies.

3.4 Metaheuristic Optimization Algorithms

The increasing complexity of deep learning models employed in economic and financial analytics has intensified the need for robust and efficient optimization strategies. In customer-level financial prediction tasks, model performance is highly sensitive to hyperparameter configurations, which govern learning dynamics, model capacity, and regularization behavior. Manual or grid-based tuning approaches are often infeasible in such settings due to the high dimensionality of the hyperparameter space and the substantial computational cost associated with repeated model training. Metaheuristic optimization algorithms offer a powerful alternative by providing flexible, population-based search mechanisms capable of navigating complex, non-convex optimization landscapes.

Metaheuristic algorithms are particularly well suited for financial analytics applications because they do not rely on gradient information or strong assumptions about the objective function. This characteristic is essential when optimizing deep learning models, where loss surfaces are often highly nonlinear and multimodal. By iteratively exploring and exploiting the search space, metaheuristics enable automated and adaptive tuning of model hyperparameters, thereby enhancing predictive accuracy and robustness while reducing the risk of suboptimal model configurations.

3.4.1 Role of Metaheuristics in Hyperparameter Optimization

In this study, metaheuristic optimization algorithms are employed to automate the tuning of hyperparameters for the Seq2Seq architecture as well as other baseline deep learning models. Hyperparameters such as learning rates, network depth, hidden unit dimensions, and regularization coefficients play a decisive role in determining model performance, convergence stability, and generalization capability. In the context of customer-level financial prediction, inappropriate hyperparameter settings can lead to overfitting, unstable learning behavior, or poor out-of-sample performance, all of which undermine the economic reliability of predictive models.

The use of metaheuristics enables a systematic and data-driven exploration of the hyperparameter space, allowing the optimization process to adaptively identify configurations that balance predictive accuracy and generalization performance. This balance is particularly critical in economic and financial applications, where models must not only fit historical data but also maintain stability when applied to unseen customer profiles. By incorporating generalization-oriented objective functions, the optimization process prioritizes solutions that achieve strong predictive performance without excessive complexity.

Furthermore, metaheuristic-based hyperparameter optimization supports scalability and reproducibility in applied financial analytics. Once an optimization framework is established, it can be consistently applied across different models and datasets, reducing reliance on ad hoc tuning decisions. This systematic approach enhances methodological transparency and ensures that performance improvements can be attributed to optimized model configurations rather than subjective parameter choices.

3.5 Training and Optimization Using POSC

In customer-level economic and financial analytics, the predictive accuracy and robustness of deep learning models are strongly affected by the configuration of hyperparameters and, in many cases, by the composition of the feature set. This sensitivity is particularly important in supervised financial prediction settings, where the objective is to construct models that generalize reliably across heterogeneous customer profiles rather than merely fitting the training data. To address these challenges, the present study adopts a population-based metaheuristic strategy for automated optimization. Specifically, instead of relying exclusively on conventional gradient-based optimizers for model configuration search, a hybrid Puma Optimizer–Sine Cosine Optimizer (POSC) is employed to enhance convergence behavior, improve generalization, and systematically identify high-quality configurations in a non-convex search landscape.

The Puma Optimizer (PO) [29] is inspired by the adaptive hunting behavior of pumas and is designed to alternate between exploration and exploitation. Exploration corresponds to searching diverse regions of the solution space to avoid premature convergence, whereas exploitation corresponds to intensifying the search around promising candidate solutions to refine them. This balance is essential in financial prediction tasks, where the optimization landscape is typically rugged due to nonlinear model structures and the interaction effects among demographic, behavioral, and financial features. The Sine Cosine Optimization component (SC) [30] complements PO by applying sinusoidal position update dynamics that enhance diversity and provide a controlled mechanism for moving candidate solutions toward the current best solution. In combination, POSC provides an adaptive hybrid mechanism that supports both global exploration and local refinement.

Optimization variables and solution representation. Let a candidate solution be represented by a position vector P that encodes the decision variables subject to optimization. In the present framework, P is defined to include the model configuration parameters (e.g., hyperparameters of the baseline deep learning model under consideration) and, when feature selection is activated, a binary mask indicating the selected subset of input variables. Accordingly, each candidate solution can be expressed as

$$P = \big[\boldsymbol{\theta}, \mathbf{z}\big], \tag{1}$$

where $\boldsymbol{\theta}$ denotes the continuous and/or discrete hyperparameter set, and $\mathbf{z} \in \{0,1\}^d$ denotes a binary feature-selection vector of length d, with $z_j = 1$ indicating that the j-th feature is retained.

Fitness function for financial prediction. Because the target variable is a continuous financial indicator (e.g., credit score or estimated savings), the optimization objective is defined in terms of a regression loss rather than a classification loss. Let $\hat{y}_i$ denote the predicted value for observation i and y_i the observed target value. The mean squared error (MSE) on the validation set is defined as

$$\text{MSE} = \frac{1}{m}\sum_{i=1}^{m}\big(y_i - \hat{y}_i\big)^2, \tag{2}$$

where m denotes the number of validation samples. Since financial prediction studies often report error in the original scale of the target variable, the root mean squared error (RMSE) is also considered:

$$\text{RMSE} = \sqrt{\text{MSE}}. \tag{3}$$

In this study, the fitness function is defined to minimize the validation error of the model produced by a candidate solution, which can be expressed generically as

$$\min_{P}\ \mathcal{F}(P), \tag{4}$$

where $\mathcal{F}(P)$ denotes the chosen validation-based objective (e.g., RMSE), computed after training the model with configuration $\boldsymbol{\theta}$ using the feature subset indicated by $\mathbf{z}$.

Within PO, the decision to prioritize exploration or exploitation is guided by scores that quantify the relative desirability of the two search modes. These scores are defined as

$$Score_{\text{Explore}} = (PF_1 \cdot f1_{\text{Explore}}) + (PF_2 \cdot f2_{\text{Explore}}), \tag{5}$$

$$Score_{\text{Exploit}} = (PF_1 \cdot f1_{\text{Exploit}}) + (PF_2 \cdot f2_{\text{Exploit}}), \tag{6}$$

where PF_1 and PF_2 are weighting factors, and $f1(\cdot)$ and $f2(\cdot)$ represent exploration and exploitation cost functions, respectively. By comparing $Score_{\text{Explore}}$ and $Score_{\text{Exploit}}$, the algorithm adaptively selects whether to diversify the search to new regions or intensify around locally promising areas.

After the PO-driven update, the SC component refines candidate positions through sinusoidal motion toward the global best solution. Let $P(t)$ denote the position vector at iteration t, and let $S^*(t)$ denote the best solution found up to iteration t. The SC update is defined as

$$P(t+1) = \begin{cases} P(t) + r_5 \sin(r_6)\,\big|r_7 S^*(t) - S(t)\big|, & \text{if } r_4 < 0.5, \\ P(t) + r_5 \cos(r_6)\,\big|r_7 S^*(t) - S(t)\big|, & \text{if } r_4 \geq 0.5, \end{cases} \tag{7}$$

where $r_4, r_5, r_6,$ and r_7 are random numbers that introduce stochasticity and maintain diversity. This refinement step is particularly valuable when the optimization landscape contains multiple local minima, a common feature in deep learning hyperparameter optimization for economic prediction tasks.

3.5.1 State-of-the-Art Metaheuristic Models

To ensure a comprehensive and representative optimization framework, this study employs a diverse set of state-of-the-art metaheuristic algorithms that have demonstrated effectiveness across a wide range of optimization problems. The selected algorithms represent different evolutionary and swarm-based paradigms, allowing for complementary exploration and exploitation behaviors within the hyperparameter search space.

The Genetic Algorithm (GA) is included as a classical evolutionary optimization technique inspired by natural selection and genetic inheritance. GA employs operators such as selection, crossover, and mutation to iteratively evolve candidate solutions, making it well suited for exploring complex and discrete hyperparameter spaces commonly encountered in deep learning optimization.

The Bat Algorithm (BA) is adopted as a bio-inspired metaheuristic that simulates the echolocation behavior of bats. Its adaptive frequency and loudness mechanisms enable dynamic adjustment between global exploration and local exploitation, which is particularly beneficial for fine-tuning deep learning hyperparameters in financial prediction tasks.

Differential Evolution (DE) is incorporated due to its strong performance in continuous optimization problems. DE relies on vector-based mutation and recombination strategies that promote efficient exploration of the search space, making it suitable for optimizing continuous-valued hyperparameters such as learning rates and regularization coefficients.

The Whale-based optimization approach, denoted as WAO, is employed to capture swarm intelligence dynamics inspired by the cooperative foraging behavior of whales. Its balance between exploration and exploitation allows for effective navigation of complex optimization landscapes, which is essential when tuning high-capacity deep learning models.

Multiverse Optimization (MVO) is included as a physics-inspired metaheuristic that models candidate solutions as universes connected through probabilistic exchange mechanisms. MVO's ability to balance diversification and intensification supports efficient search in high-dimensional hyperparameter spaces relevant to financial analytics.

Particle Swarm Optimization (PSO) is selected due to its simplicity, fast convergence, and proven effectiveness in continuous optimization settings. By leveraging collective learning and information sharing among particles, PSO provides a robust mechanism for identifying high-quality hyperparameter configurations with relatively low computational overhead.

Stochastic Fractal Search (SFS) is incorporated as an advanced metaheuristic that combines stochastic diffusion processes with fractal-based exploration. Its strong local search capabilities complement the global exploration behavior of other algorithms, contributing to a well-rounded optimization framework.

Together, these metaheuristic algorithms form a comprehensive optimization toolkit for hyperparameter tuning in deep learning–based financial prediction models. Their collective use enables a rigorous examination of how different optimization strategies influence model performance, stability, and generalization in customer-level economic and financial analytics.

3.6 Evaluation Metrics

The performance of the proposed deep learning and optimization frameworks is assessed using a comprehensive set of regression-based evaluation metrics that are widely adopted in economic, financial, and forecasting studies. The use of multiple metrics is essential in customer-level financial prediction, as no single indicator can fully capture all aspects of model performance. While error-based metrics quantify the magnitude of prediction deviations, correlation- and efficiency-based measures provide complementary insights into goodness of fit, bias behavior, and predictive reliability.

In financial analytics, accurate error quantification is crucial because even small deviations can translate into substantial economic costs or misinformed decision-making. Accordingly, absolute and squared error measures are employed to capture both average deviations and the penalization of large errors. In addition, bias-sensitive and relative error metrics are included to assess systematic over- or underestimation tendencies, which are particularly relevant in credit-related and savings-related prediction tasks. Finally, efficiency and agreement indices are used to evaluate how well the predictive models reproduce observed variability and overall patterns in the target variable.

Table 1 summarizes the evaluation metrics used in this study along with their mathematical definitions. Let y_i denote the observed financial target value for the i-th observation, $\hat{y}_i$ the corresponding predicted value, $\bar{y}$ the mean of the observed values, and n the total number of samples.

Table 1: Regression evaluation metrics used for model performance assessment

Metric	**Mathematical Definition**
Mean Squared Error (MSE)	$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$
Root Mean Squared Error (RMSE)	$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$
Mean Absolute Error (MAE)	$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\lvert y_i - \hat{y}_i\rvert$
Mean Bias Error (MBE)	$\text{MBE} = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)$
Correlation Coefficient (r)	$r = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}\sqrt{\sum_{i=1}^{n}(\hat{y}_i - \bar{\hat{y}})^2}}$
Coefficient of Determination (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)	$\text{RRMSE} = \frac{\text{RMSE}}{\bar{y}} \times 100$
Nash–Sutcliffe Efficiency (NSE)	$\text{NSE} = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$
Willmott's Index of Agreement (WI)	$\text{WI} = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}\left(\lvert\hat{y}_i - \bar{y}\rvert + \lvert y_i - \bar{y}\rvert\right)^2}$

The combined use of these metrics ensures a rigorous and multidimensional evaluation of predictive performance. Error-based indicators such as MSE, RMSE, and MAE quantify absolute accuracy, while MBE reveals systematic prediction bias. The correlation coefficient and coefficient of determination assess the strength and explanatory power of the relationship between observed and predicted values. Relative measures such as RRMSE provide scale-independent assessment, which is particularly relevant when comparing financial indicators across different magnitudes. Finally, NSE and Willmott's Index offer efficiency- and agreement-based perspectives that are widely used in applied forecasting and economic modeling to evaluate overall predictive skill.

4 Experimental Results

4.1 Baseline Model Performance (Before Optimization)

This subsection provides a detailed comparative analysis of the baseline performance of the deep learning models prior to the application of any feature selection or metaheuristic-based optimization techniques. Establishing baseline results is a fundamental step in empirical economic and financial modeling, as it allows researchers to assess the intrinsic predictive capability of each model architecture when trained using standard configurations and the complete set of available features. In customer-level financial analytics, such baseline evaluations are particularly important because they reveal how well different models capture heterogeneous consumer behavior and financial dynamics without the aid of specialized optimization procedures.

The baseline models evaluated in this study include Seq2Seq, N-HITS, Informer, NODE, and GRU. These models were selected to represent a diverse range of deep learning paradigms, including sequence-to-sequence learning, hierarchical forecasting structures, attention-based architectures, continuous-time neural modeling, and gated recurrent networks. By examining these architectures under identical data and evaluation conditions, the analysis ensures a fair and consistent comparison of their performance in predicting financially relevant customer-level indicators.

Table 2 summarizes the baseline performance of all models using the regression metrics introduced in Section 3.6. The reported metrics collectively capture multiple dimensions of predictive quality, including absolute and squared prediction errors, systematic bias, correlation strength, explanatory power, and efficiency. This multidimensional evaluation is essential in financial prediction tasks, where accuracy, stability, and interpretability all play critical roles in supporting economic decision-making.

Table 2: Baseline performance of deep learning models before optimization

Model	**MSE**	**RMSE**	**MAE**	**MBE**	r	R^2	**RRMSE**	**NSE**	**WI**
Seq2Seq	0.0065	0.0809	0.0441	0.0295	0.898	0.892	1.39	0.890	0.892
N-HITS	0.0091	0.0956	0.0552	0.0404	0.886	0.881	1.64	0.879	0.884
Informer	0.0131	0.1145	0.0655	0.0493	0.875	0.870	1.90	0.868	0.874
NODE	0.0179	0.1338	0.0766	0.0581	0.865	0.859	2.19	0.856	0.864
GRU	0.0231	0.1520	0.0876	0.0658	0.852	0.845	2.44	0.844	0.852

The results reported in Table 2 indicate notable differences in predictive accuracy across the baseline models. Among the evaluated architectures, Seq2Seq consistently exhibits the strongest performance across nearly all metrics. Its lower MSE, RMSE, and MAE values suggest a superior ability to minimize both average and extreme prediction errors, while its relatively low MBE indicates reduced systematic bias. Furthermore, the higher values of the correlation coefficient, coefficient of determination, Nash–Sutcliffe efficiency, and Willmott's Index reflect strong agreement between observed and predicted financial values, highlighting the effectiveness of Seq2Seq in capturing complex nonlinear relationships in customer-level financial data.

N-HITS demonstrates the second-best baseline performance, with moderate increases in error-based metrics relative to Seq2Seq. Its hierarchical modeling structure appears to capture relevant financial patterns, although with reduced precision and slightly higher bias. Informer follows closely, benefiting from its attention-based architecture, which allows it to focus selectively on informative features. However, under baseline configurations, its performance remains inferior to Seq2Seq and N-HITS, suggesting that attention mechanisms alone may be insufficient to fully exploit the structure of enriched customer financial data without further optimization.

The NODE and GRU models exhibit comparatively weaker baseline performance, as evidenced by higher error values and lower efficiency scores. While these models are capable of modeling nonlinear relationships and sequential dependencies, their baseline configurations appear less effective in capturing the full complexity of the financial prediction task. The higher RRMSE values observed for these models indicate reduced relative accuracy, which may limit their practical applicability in sensitive financial decision-making contexts without additional refinement.

From an economic and financial perspective, the observed baseline performance differences underscore the importance of model architecture selection in customer analytics. The results suggest that certain deep learning designs are inherently better suited to capturing the joint influence of demographic, behavioral, and financial variables. At the same time, the presence of non-negligible prediction errors and systematic bias across all baseline models highlights the limitations of unoptimized configurations. These findings provide a strong empirical motivation for the application of feature selection and metaheuristic-based hyperparameter optimization in subsequent sections, where the goal is to enhance predictive accuracy, reduce bias, and improve generalization performance in customer-level financial analytics.

To provide a comprehensive evaluation of model performance stability and accuracy, we analyze the empirical distributions of multiple error- and skill-based metrics using combined density and kernel density estimation (KDE) visualizations. This approach enables simultaneous assessment of central tendency, dispersion, and distributional shape, offering richer insight than point estimates alone. As shown in Figure 4, the density–KDE plots for mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), and mean bias error (MBE) highlight the variability and skewness associated with prediction errors. In parallel, correlation-based and efficiency metrics—including the Pearson correlation coefficient (r), coefficient of determination (R^2), relative root mean squared error (RRMSE), Nash–Sutcliffe efficiency (NSE), and Willmott's index (WI)—exhibit comparatively tighter distributions, reflecting consistent predictive skill across experimental runs. Collectively, these visual diagnostics facilitate a robust comparison of model reliability and generalization behavior, thereby supporting informed interpretation of predictive performance.

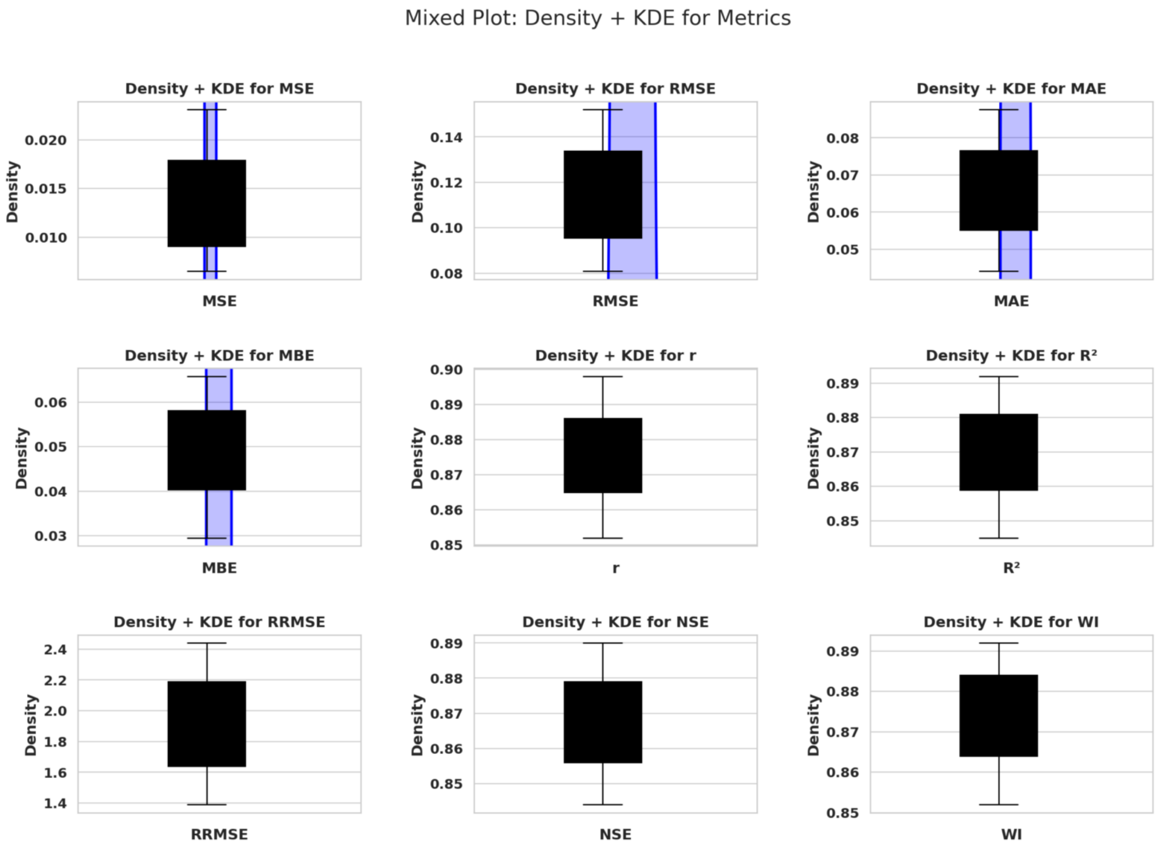

Figure 4: Combined density and kernel density estimation plots for error-based and skill-based performance metrics, illustrating distributional characteristics across multiple evaluation criteria.

To enhance the interpretability of model evaluation results, we employ an integrated visualization framework that combines swarm plots, violin plots, and boxplots for a comprehensive set of performance metrics. This multi-layered approach enables simultaneous inspection of individual observations, distributional density, and robust summary statistics, thereby providing a detailed perspective on both variability and central tendency. As illustrated in Figure 5, the error-based metrics—mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), mean bias error (MBE), and relative root mean squared error (RRMSE)—exhibit moderate dispersion with well-defined medians, indicating stable predictive accuracy across experimental runs. In contrast, skill-based indicators, including the Pearson correlation coefficient (r), coefficient of determination (R^2), Nash–Sutcliffe efficiency (NSE), and Willmott's index (WI), display comparatively concentrated distributions at higher values, reflecting consistent model skill and strong agreement between observed and predicted values. Overall, this figure facilitates a holistic assessment of model robustness and performance consistency by integrating complementary visualization techniques within a unified analytical view.

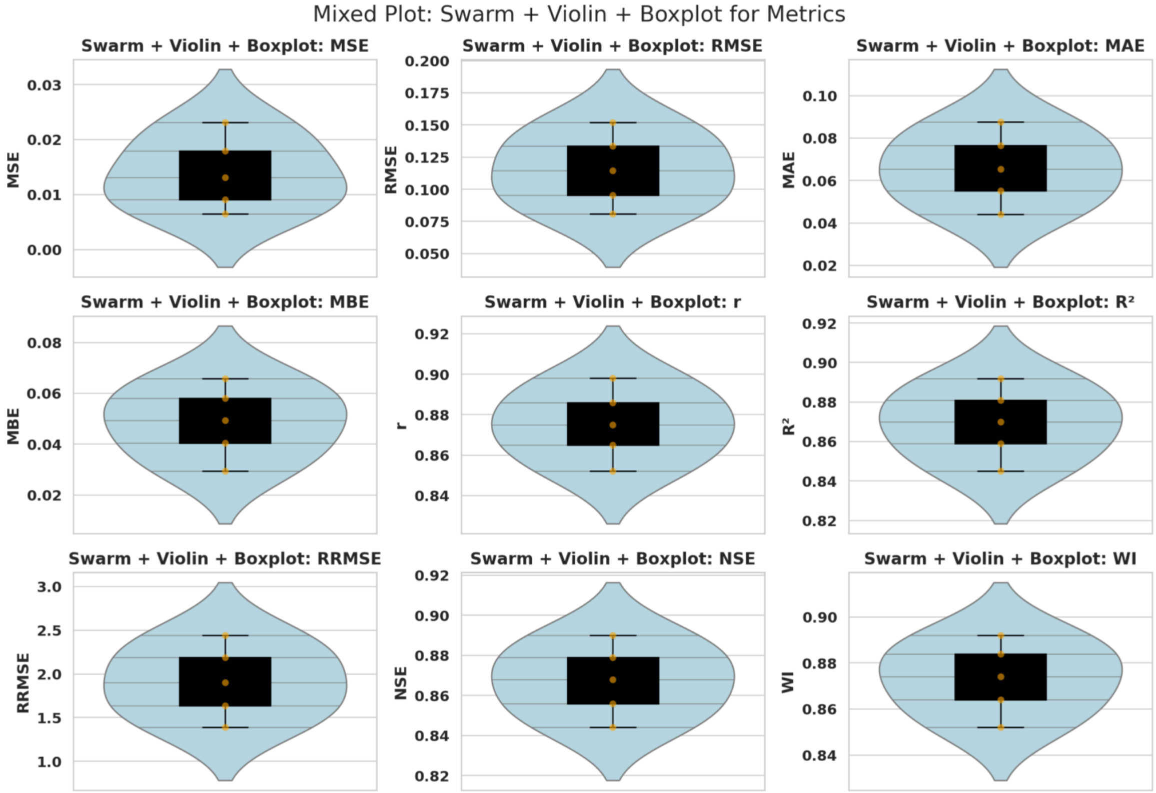

Figure 5: Combined swarm, violin, and boxplot representations for multiple error-based and skill-based performance metrics, illustrating distributional structure, central tendency, and individual observations.

To evaluate and contrast the predictive accuracy of the considered forecasting models, we conduct a comparative analysis using widely adopted error-based performance metrics. Specifically, mean squared error (MSE), root mean squared error (RMSE), and mean absolute error (MAE) are employed to quantify both the magnitude and consistency of prediction errors across different model architectures. As illustrated in Figure 6, the bar chart facilitates a direct comparison of model performance by highlighting relative differences in error levels. Lower values of these metrics indicate superior predictive accuracy, enabling clear identification of models that achieve more reliable forecasts. The observed variation across models reflects differences in their capacity to capture temporal dependencies and nonlinear patterns in the data, thereby providing an empirical basis for model selection and further optimization.

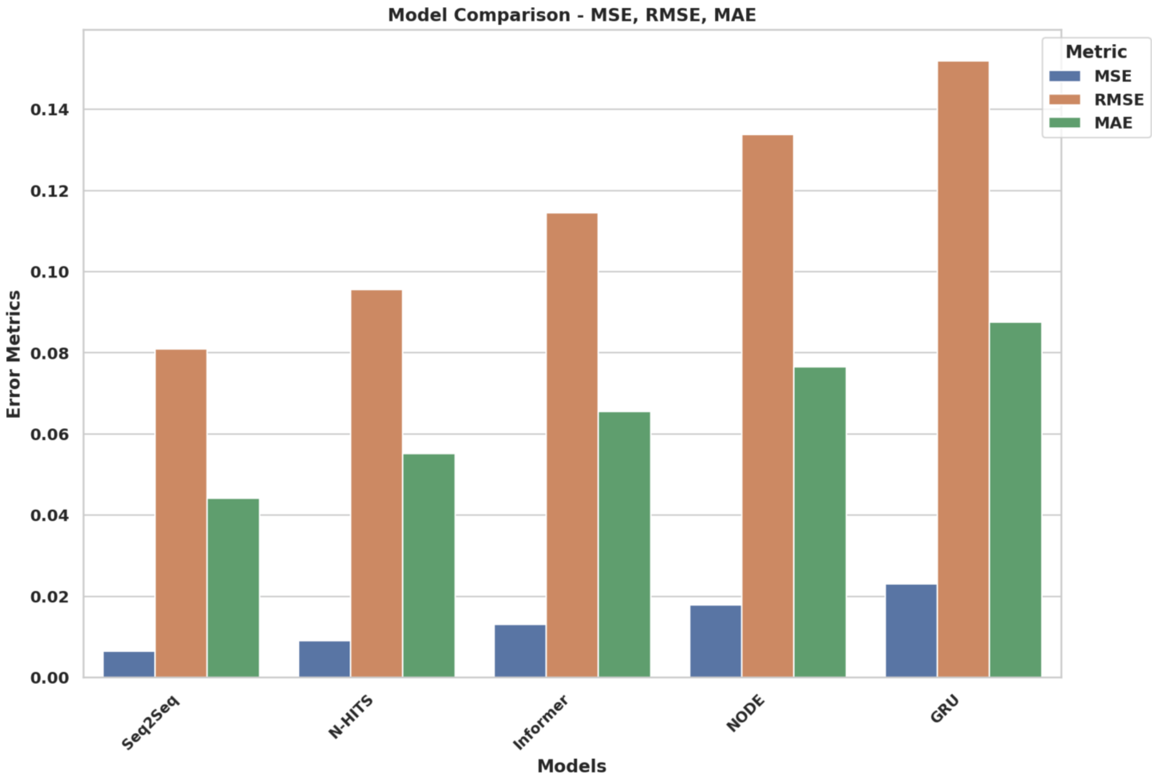

Figure 6: Comparison of forecasting models based on mean squared error, root mean squared error, and mean absolute error, illustrating relative predictive accuracy across model architectures.

To synthesize model evaluation results while explicitly accounting for variability, we present a comparative visualization of mean performance metrics accompanied by error bars. This representation enables simultaneous assessment of average model behavior and the associated uncertainty across multiple evaluation criteria. As illustrated in Figure 7, the plotted metrics include error-based indicators—mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), and mean bias error (MBE)—as well as skill-based measures such as the Pearson correlation coefficient (r), coefficient of determination (R^2), relative root mean squared error (RRMSE), Nash–Sutcliffe efficiency (NSE), and Willmott's index (WI). The inclusion of error bars highlights the dispersion around each mean value, providing insight into the robustness and stability of model performance. This visualization supports hints on trade-offs between accuracy and reliability, thereby facilitating a more informed interpretation of overall predictive quality.

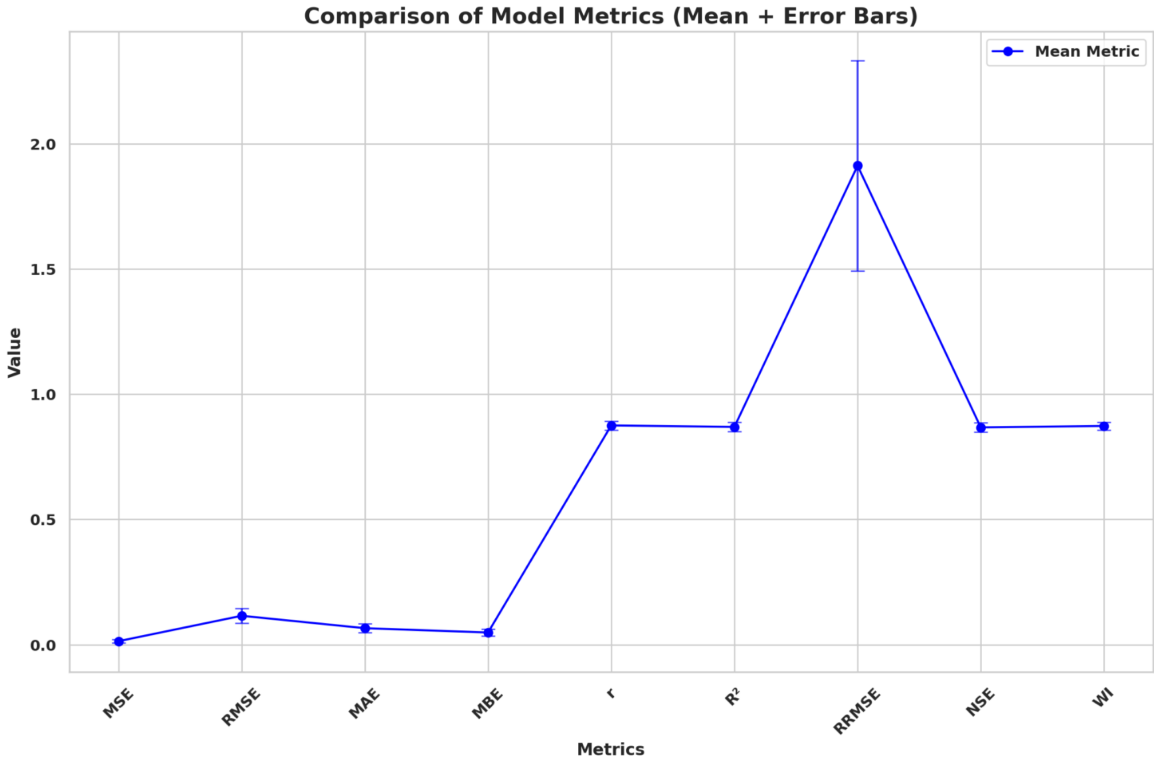

Figure 7: Comparison of mean performance metrics with corresponding error bars, illustrating average predictive behavior and associated uncertainty across multiple evaluation criteria.

4.2 Optimized Model Analysis

In this section, the performance of the optimized models is analyzed after applying metaheuristic-based optimization strategies to the Seq2Seq architecture. The optimization process focuses on improving model generalization, convergence stability, and predictive accuracy by systematically tuning hyperparameters using population-based metaheuristic algorithms. By fixing the underlying deep learning architecture and varying only the optimization strategy, the analysis isolates the contribution of each metaheuristic algorithm to performance enhancement in customer-level financial prediction.

The optimization techniques considered in this study include Puma Optimizer (PO), Genetic Algorithm (GA), Bat Algorithm (BA), Differential Evolution (DE), Whale-based Optimization (WAO), Multiverse Optimization (MVO), Particle Swarm Optimization (PSO), and Stochastic Fractal Search (SFS), each combined with the Seq2Seq model. These algorithms represent diverse optimization philosophies, encompassing evolutionary mechanisms, swarm intelligence, and stochastic diffusion processes. Their application aims to identify high-quality hyperparameter configurations that balance predictive accuracy and generalization in a complex, non-convex optimization landscape typical of financial analytics.

The predictive performance of the optimized models is evaluated using the regression metrics defined in Section 3.6. These metrics collectively assess absolute error magnitude, systematic bias, correlation strength, explanatory power, and efficiency-based agreement between observed and predicted financial values. Table 3 presents the detailed performance comparison of the optimized Seq2Seq models under different metaheuristic optimization strategies.

Table 3: Performance comparison of optimized Seq2Seq models using different metaheuristic algorithms

Model	MSE	RMSE	MAE	MBE	r	R^2	RRMSE	NSE	WI
PO + Seq2Seq	2.05E-05	4.52E-03	2.05E-04	5.40E-05	0.987	0.983	0.059	0.986	0.988
GA + Seq2Seq	3.12E-05	5.58E-03	2.60E-04	6.80E-05	0.979	0.975	0.081	0.978	0.981
BA + Seq2Seq	4.18E-05	6.50E-03	3.00E-04	8.20E-05	0.973	0.970	0.098	0.972	0.977
DE + Seq2Seq	5.42E-05	7.36E-03	3.36E-04	9.60E-05	0.966	0.963	0.128	0.965	0.973
WAO + Seq2Seq	6.98E-05	8.38E-03	3.76E-04	1.13E-04	0.959	0.957	0.178	0.958	0.968
MVO + Seq2Seq	8.88E-05	9.44E-03	4.08E-04	1.28E-04	0.952	0.950	0.228	0.950	0.962
PSO + Seq2Seq	1.08E-04	1.04E-02	4.36E-04	1.43E-04	0.946	0.944	0.278	0.944	0.957
SFS + Seq2Seq	1.33E-04	1.15E-02	4.63E-04	1.58E-04	0.939	0.937	0.408	0.937	0.952

The results in Table 3 indicate that the PO + Seq2Seq configuration achieves the best overall performance among all optimized models. This model records the lowest MSE, RMSE, and MAE values, demonstrating a substantial reduction in both average and extreme prediction errors. The very small MBE further indicates negligible systematic bias, which is particularly important in financial applications where persistent bias can distort economic decisions. The high correlation coefficient and coefficient of determination confirm that PO + Seq2Seq captures most of the variability in the financial target variable, while the high NSE and WI values indicate excellent agreement between predicted and observed values.

The GA + Seq2Seq and BA + Seq2Seq configurations also exhibit strong predictive performance, ranking immediately after PO + Seq2Seq. Although their error metrics are slightly higher, both models maintain high correlation and efficiency scores, suggesting reliable and stable predictive behavior. These results indicate that evolutionary and bio-inspired optimization strategies are effective in enhancing the learning capability of Seq2Seq for customer-level financial analytics, albeit with varying levels of efficiency.

The DE + Seq2Seq and WAO + Seq2Seq models demonstrate moderate performance improvements. While they clearly outperform the unoptimized baseline models, their higher error and bias values relative to the top-performing configurations indicate reduced precision. Nevertheless, their relatively strong correlation and efficiency metrics suggest that these optimizers remain viable options, particularly in scenarios where different exploration–exploitation trade-offs are desired.

The MVO + Seq2Seq, PSO + Seq2Seq, and SFS + Seq2Seq configurations yield comparatively lower performance among the optimized models. Although they still provide notable improvements over baseline results, their higher RRMSE values and lower efficiency indices indicate weaker relative accuracy and agreement. These findings highlight that the effectiveness of metaheuristic optimization is strongly dependent on the compatibility between the optimizer's search dynamics and the underlying deep learning architecture.

Overall, the optimized model analysis demonstrates that metaheuristic-based hyperparameter optimization can substantially enhance the predictive accuracy, robustness, and generalization capability of deep learning models in customer-level financial prediction tasks. The results clearly show that coupling Seq2Seq with an appropriate optimizer, particularly PO, leads to superior performance across all evaluation metrics. At the same time, the variability in performance across optimization strategies underscores the importance of systematic comparative evaluation when designing optimization-driven financial analytics frameworks.

To explore structural similarities among the hybrid forecasting models and to identify groups with comparable performance characteristics, we apply hierarchical clustering based on their aggregated evaluation metrics. This unsupervised analysis provides insight into how different optimization strategies interact with the Seq2Seq architecture, revealing latent relationships that are not immediately evident from individual performance scores. As illustrated in Figure 8, the resulting dendrogram organizes the models according to their pairwise distances, where shorter linkage distances indicate stronger similarity in predictive behavior. Distinct clusters emerge, reflecting shared error profiles and skill levels among certain metaheuristic-enhanced models, while larger linkage heights highlight models with more divergent performance patterns. Overall, this hierarchical representation supports a deeper understanding of model grouping and redundancy, offering guidance for informed model selection and ensemble design.

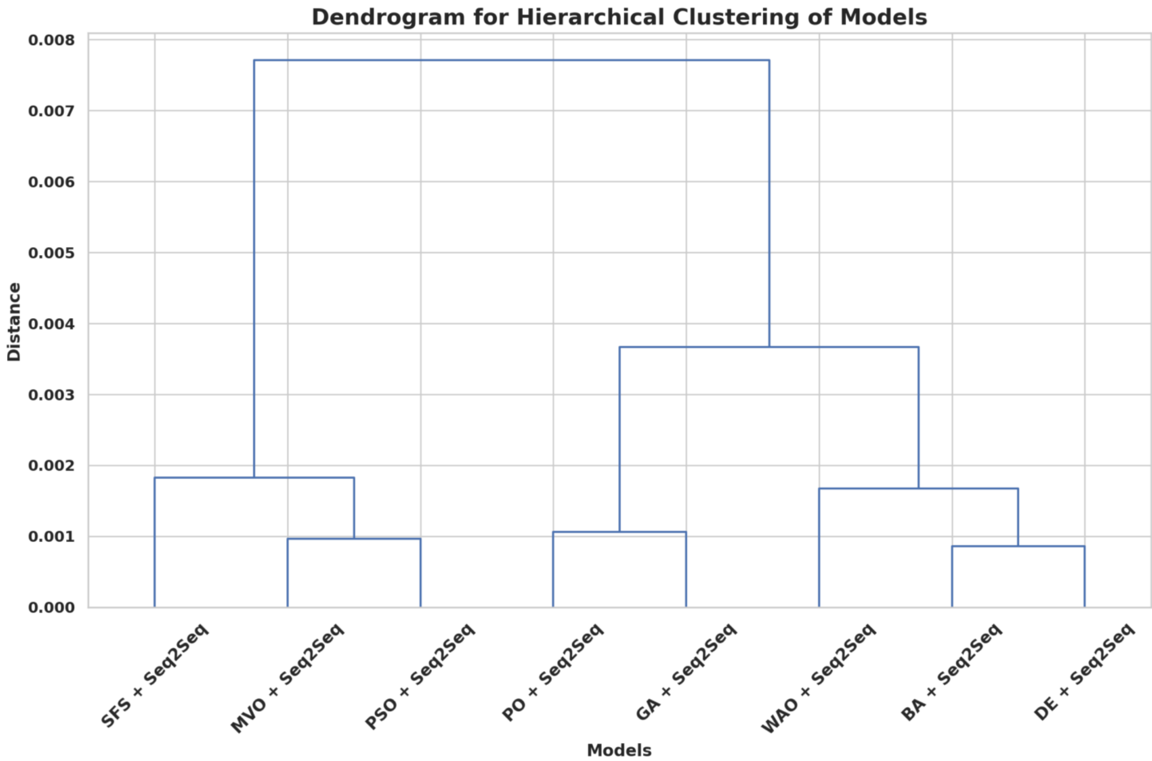

Figure 8: Dendrogram illustrating hierarchical clustering of hybrid Seq2Seq-based models, where linkage distances reflect similarities in aggregated performance metrics.

To facilitate a holistic comparison of the hybrid forecasting models across multiple and potentially competing performance criteria, we employ a radar chart representation. This visualization is particularly effective for simultaneously assessing error-based, correlation-based, and efficiency-based metrics within a unified framework. As shown in Figure 9, each axis corresponds to a normalized evaluation metric hinting at predictive accuracy and consistency, including mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), mean bias error (MBE), relative root mean squared error (RRMSE), Pearson correlation coefficient (r), and coefficient of determination (R^2). The enclosed area formed by each model reflects its overall performance profile, where larger and more balanced shapes indicate stronger and more stable predictive behavior across metrics. This radar-based comparison enables intuitive identification of trade-offs among models and highlights those achieving superior multi-metric performance.

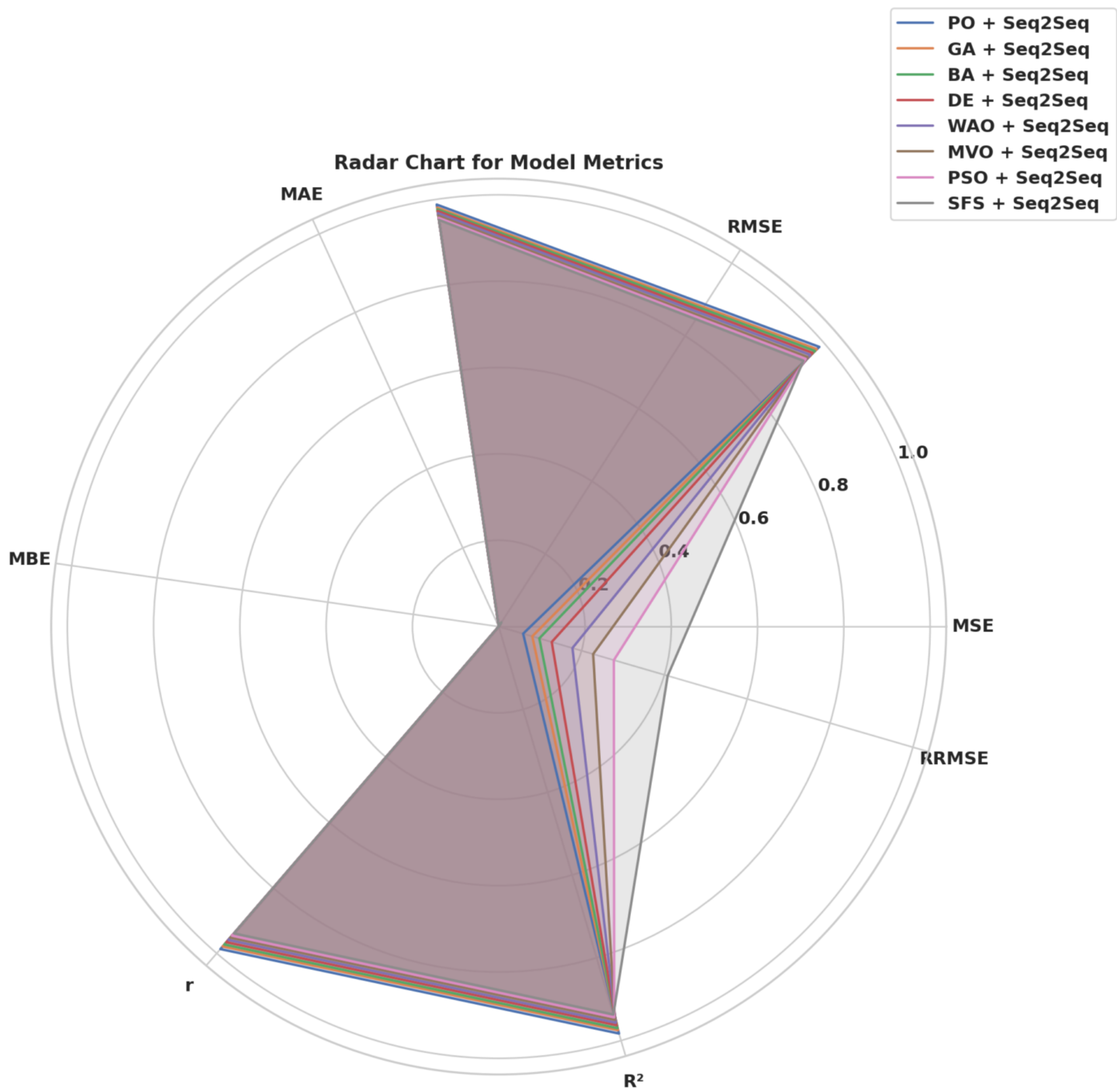

Figure 9: Radar chart illustrating normalized performance metrics for hybrid Seq2Seq-based models, enabling multi-criteria comparison of predictive accuracy and consistency.

To enable a high-dimensional comparison of hybrid forecasting models across multiple evaluation criteria, we adopt a parallel coordinates visualization. This technique is particularly well suited for identifying performance trade-offs and consistency patterns when models are assessed using heterogeneous metrics. As illustrated in Figure 10, each polyline represents a hybrid Seq2Seq-based model, while each vertical axis corresponds to a normalized performance metric, including mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), mean bias error (MBE), Pearson correlation coefficient (r), coefficient of determination (R^2), relative root mean squared error (RRMSE), Nash–Sutcliffe efficiency (NSE), and Willmott's index (WI). The convergence and divergence of lines across axes reveal similarities and differences in model behavior, with tightly clustered trajectories indicating consistent performance and pronounced deviations highlighting trade-offs among error-based and skill-based measures. Overall, this visualization provides an integrated perspective on model robustness and comparative effectiveness across the full spectrum of evaluation metrics.

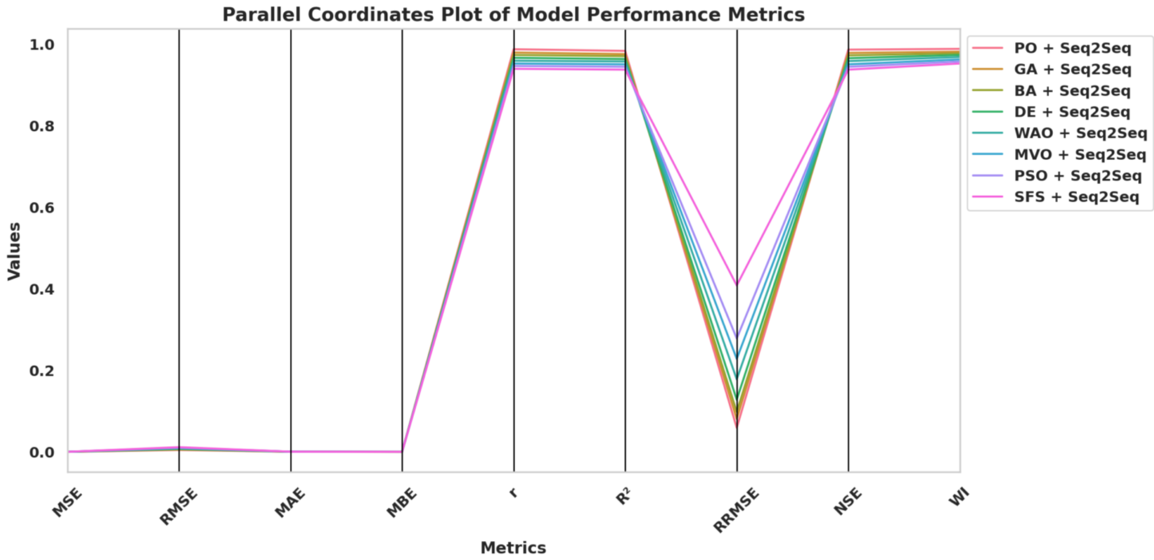

Figure 10: Parallel coordinates plot illustrating normalized performance metrics for hybrid Seq2Seq-based models, enabling simultaneous multi-criteria comparison.

To provide a structured and metric-specific comparison of model performance, we present a facet grid visualization in which each panel corresponds to a distinct evaluation criterion. This approach allows individual metrics to be examined independently while preserving a consistent comparison across models. As illustrated in Figure 11, error-based measures—including mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), mean bias error (MBE), and relative root mean squared error (RRMSE)—exhibit clear monotonic variations across models, highlighting differences in predictive accuracy and bias. In contrast, skill-based indicators such as the Pearson correlation coefficient (r), coefficient of determination (R^2), Nash–Sutcliffe efficiency (NSE), and Willmott's index (WI) remain comparatively stable and concentrated at higher values, reflecting consistent explanatory power and agreement between predicted and observed values. This facet-based representation facilitates focused interpretation of each performance dimension and supports transparent, side-by-side evaluation of competing models.

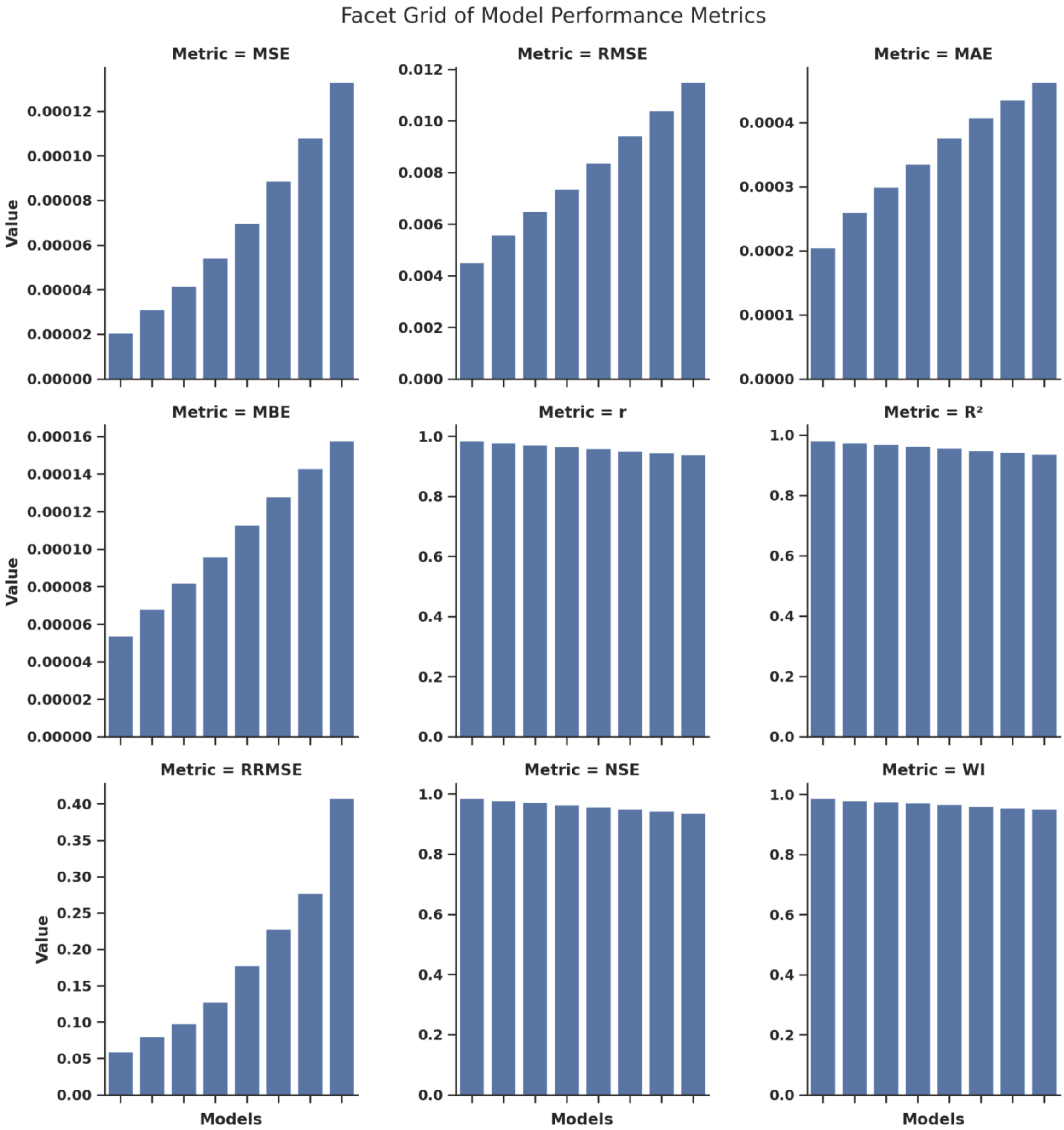

Figure 11: Facet grid visualization of model performance metrics, where each panel represents a distinct error-based or skill-based evaluation criterion across different models.

5 Conclusion and Future Work

This study has presented a comprehensive machine learning–based framework for customer-level financial prediction, with a particular focus on integrating deep learning architectures and metaheuristic optimization techniques within an economic and financial analytics context. By leveraging an enhanced customer dataset that combines demographic characteristics, behavioral indicators, and financially relevant attributes, the proposed framework addresses key challenges associated with high-dimensionality, nonlinear relationships, and model sensitivity in modern retail and consumer finance applications.

The empirical analysis demonstrates that advanced deep learning models are capable of capturing complex interactions among income, spending behavior, savings tendencies, and customer engagement variables. However, the baseline evaluation also highlights that model architecture alone is insufficient to achieve optimal predictive performance. Without appropriate optimization, even sophisticated models exhibit non-negligible prediction errors and systematic bias, which may limit their reliability in real-world financial decision-making. This finding underscores the importance of coupling deep learning with robust optimization strategies when dealing with heterogeneous customer data.

The incorporation of metaheuristic algorithms for hyperparameter optimization significantly enhances model performance and generalization capability. In particular, the results show that carefully designed population-based optimization strategies can effectively navigate complex, non-convex search spaces and identify high-quality configurations that substantially reduce prediction error and bias. From an economic perspective, these improvements translate into more reliable estimates of financially relevant indicators, which are essential for applications such as customer valuation, credit-related assessment, and targeted marketing strategy.

Overall, the proposed framework contributes to the growing literature on data-driven economic and financial analytics by demonstrating the effectiveness of combining deep learning with metaheuristic optimization in customer-level prediction tasks. The findings highlight that optimization-driven modeling approaches can provide substantial gains in accuracy and robustness, thereby supporting more informed and efficient decision-making in retail economics and consumer finance.

While the results obtained in this study are encouraging, several directions for future research can be identified. First, future work may explore the integration of additional economic variables, such as macroeconomic indicators or regional economic characteristics, to assess how broader economic conditions influence customer-level financial behavior. Incorporating such contextual information could further enhance predictive accuracy and improve the interpretability of model outcomes from a policy and strategic planning perspective.

Second, extending the proposed framework to dynamic or longitudinal datasets represents a promising avenue for future research. Customer behavior and financial status evolve over time, and modeling these temporal dynamics explicitly may yield deeper insights into consumption smoothing, lifecycle effects, and long-term customer value. The integration of time-aware learning architectures and adaptive optimization strategies could support more accurate forecasting in evolving economic environments.

Third, future studies may investigate hybrid optimization frameworks that combine multiple metaheuristic strategies or integrate metaheuristics with gradient-based fine-tuning. Such hybrid approaches could further improve convergence speed and solution quality, particularly for large-scale financial datasets. Additionally, analyzing the computational trade-offs associated with different optimization strategies would be valuable for practical deployment in real-world financial systems.

Finally, the practical implementation of the proposed framework in decision-support systems for retail and financial institutions warrants further investigation. Issues related to scalability, real-time prediction, and interpretability are critical for operational adoption. Future research could focus on embedding optimized predictive models into intelligent financial platforms that support dynamic pricing, personalized financial products, and risk-aware customer management, thereby strengthening the link between advanced analytics and economically meaningful outcomes.

Data Availability

The dataset used in this study is publicly available on Kaggle at https://www.kaggle.com/datasets/dhrubangtalukdar/goldman-sachs-all-time-stock-data.

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] Y. Kim and R. J. Kauffman, "Consumer broadband choices: Smartphone and digital entertainment growth in e-commerce," *International Journal of Electronic Commerce*, vol. 29, no. 3, pp. 366–400, 2025. DOI: 10.1080/10864415.2025.2515727. eprint: https://doi.org/10.1080/10864415.2025.2515727. [Online]. Available: https://doi.org/10.1080/10864415.2025.2515727.

[2] Z. Shu, M. Llorens-Marin, R. A. Carrasco, and M. S. Romero, "Customer electronic word of mouth management strategies based on computing with words: The case of spanish luxury hotel reviews on tripadvisor," *Electronics*, vol. 14, no. 2, 2025, ISSN: 2079-9292. DOI: 10.3390/electronics14020325. [Online]. Available: https://www.mdpi.com/2079-9292/14/2/325.

[3] S. B. Friend, S. M. Mangus, E. B. Pullins, L. Davis, and C. M. Gilstrap, "Conceptualizing an integrative typology of sales enablement strategy," *Journal of Personal Selling & Sales Management*, vol. 45, no. 1, pp. 64–81, 2025. DOI: 10.1080/08853134.2024.2340033. eprint: https://doi.org/10.1080/08853134.2024.2340033. [Online]. Available: https://doi.org/10.1080/08853134.2024.2340033.

[4] H. Mo and S. Ouyang, "(generative) ai in financial economics," *Journal of Chinese Economic and Business Studies*, vol. 23, no. 4, pp. 509–587, 2025. DOI: https://doi.org/10.1080/14765284.2025.2569006.

[5] A. Cascavilla, R. Caferra, A. Morone, and P. Morone, "Experimental evidence on consumers' willingness to pay in the sustainable fashion industry," *Scientific Reports*, vol. 15, no. 1, p. 38 752, 2025. DOI: https://doi.org/10.1038/s41598-025-23008-9.

[6] E. Mogaji, "Market and customer insights," in *Strategic Marketing Management: Principles and Practice*, Springer, 2025, pp. 33–63. DOI: https://doi.org/10.1007/978-3-031-99367-1_3.

[7] S. P. Migkos, N. T. Giannakopoulos, and D. P. Sakas, "Impact of influencer marketing on consumer behavior and online shopping preferences," *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 20, no. 2, 2025, ISSN: 0718-1876. DOI: 10.3390/jtaer20020111. [Online]. Available: https://www.mdpi.com/0718-1876/20/2/111.

[8] D. Boldureanu, I. Gutu, and G. Boldureanu, "Understanding the dynamics of e-wom in food delivery services: A smartpls analysis of consumer acceptance," *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 20, no. 1, 2025, ISSN: 0718-1876. DOI: 10.3390/jtaer20010018. [Online]. Available: https://www.mdpi.com/0718-1876/20/1/18.

[9] C. Daniel Santos, "The tails of firm growth, granularity, and business cycles," *International Journal of Industrial Organization*, vol. 105, p. 103 258, 2026, ISSN: 0167-7187. DOI: https://doi.org/10.1016/j.ijindorg.2026.103258. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167718726000111.

[10] M. M. Naveed, T. Liu, S. Mustafa, and X. Chen, "Unraveling the nexus between competition and banking efficiency in an emerging economy: A two-stage stochastic frontier analysis framework," *Systems*, vol. 13, no. 5, 2025, ISSN: 2079-8954. DOI: 10.3390/systems13050354. [Online]. Available: https://www.mdpi.com/2079-8954/13/5/354.

[11] M. Imani, M. Joudaki, A. Beikmohammadi, and H. R. Arabnia, "Customer churn prediction: A systematic review of recent advances, trends, and challenges in machine learning and deep learning," *Machine Learning and Knowledge Extraction*, vol. 7, no. 3, 2025, ISSN: 2504-4990. DOI: 10.3390/make7030105. [Online]. Available: https://www.mdpi.com/2504-4990/7/3/105.

[12] S. Polamarasetti, M. R. Kumar Kakarala, H. Gadam, J. B. Butani, S. K. Rongali, and S. K. Prajapati, "Enhancing strategic business decisions with ai-powered forecasting models in salesforce crmt," in *2025 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC)*, 2025, pp. 1–10. DOI: 10.1109/ASSIC64892.2025.11158411.

[13] S. Kumar, D. Sharma, S. Rao, W. M. Lim, and S. K. Mangla, "Past, present, and future of sustainable finance: Insights from big data analytics through machine learning of scholarly research," *Annals of Operations Research*, vol. 345, no. 2, pp. 1061–1104, 2025. DOI: https://doi.org/10.1007/s10479-021-04410-8.

[14] A. De Maio, R. Musmanno, and F. Vocaturo, "Unbiased decision making in location-routing problems with uncertain customer demands," *Soft Computing*, vol. 27, no. 18, pp. 12 883–12 893, 2023. DOI: https://doi.org/10.1007/s00500-022-06785-7.

[15] F. Hosseinzadeh Lotfi, T. Allahviranloo, W. Pedrycz, M. Shahriari, H. Sharafi, and S. Razipour GhalehJough, "Foundations of decision," in *Fuzzy Decision Analysis: Multi Attribute Decision Making Approach*. Cham: Springer International Publishing, 2023, pp. 1–56, ISBN: 978-3-031-44742-6. DOI: 10.1007/978-3-031-44742-6_1. [Online]. Available: https://doi.org/10.1007/978-3-031-44742-6_1.

[16] S. Guercini, A. La Rocca, and I. Snehota, "Decisions when interacting in customer-supplier relationships," *Industrial Marketing Management*, vol. 105, pp. 380–387, 2022, ISSN: 0019-8501. DOI: https://doi.org/10.1016/j.indmarman.2022.06.019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0019850122001572.

[17] J. Lin, S. Yang, K. Huang, K. Wang, and S. Jang, "Network- and demand-driven initialization strategy for enhanced heuristic in uncapacitated facility location problem," *Mathematics*, vol. 13, no. 13, 2025, ISSN: 2227-7390. DOI: 10.3390/math13132138. [Online]. Available: https://www.mdpi.com/2227-7390/13/13/2138.

[18] H. Sargeant, "Algorithmic decision-making in financial services: Economic and normative outcomes in consumer credit," *AI and Ethics*, vol. 3, no. 4, pp. 1295–1311, 2023. DOI: https://doi.org/10.1007/s43681-022-00236-7.

[19] K.-Y. Lin, "Generative artificial intelligence–driven sustainable supply chain management: A unisone framework for smart logistics and predictive analytics under industry 5.0," *International Journal of Logistics Research and Applications*, vol. 0, no. 0, pp. 1–32, 2025. DOI: 10.1080/13675567.2025.2540855. eprint: https://doi.org/10.1080/13675567.2025.2540855. [Online]. Available: https://doi.org/10.1080/13675567.2025.2540855.

[20] R. Buijsse, M. Willemsen, and C. Snijders, "Data-driven decision-making," in *Data Science for Entrepreneurship: Principles and Methods for Data Engineering, Analytics, Entrepreneurship, and the Society*, W. Liebregts, W.-J. van den Heuvel, and A. van den Born, Eds. Cham: Springer International Publishing, 2023, pp. 239–277, ISBN: 978-3-031-19554-9. DOI: 10.1007/978-3-031-19554-9_11. [Online]. Available: https://doi.org/10.1007/978-3-031-19554-9_11.

[21] K. Tafakkori, R. Tavakkoli-Moghaddam, and A. Siadat, "Scheduling multi-configuration last-mile delivery logistics by learning from optimisation feedback and customer preferences," *International Journal of Production Research*, vol. 63, no. 21, pp. 7835–7864, 2025. DOI: 10.1080/00207543.2025.2507795. eprint: https://doi.org/10.1080/00207543.2025.2507795. [Online]. Available: https://doi.org/10.1080/00207543.2025.2507795.

[22] S. Mukherjee and B. Datta, "Counterfeit conundrum: Insights into bottom of the pyramid consumer behavior in india," *Journal of Retailing and Consumer Services*, vol. 82, p. 104 114, 2025, ISSN: 0969-6989. DOI: https://doi.org/10.1016/j.jretconser.2024.104114. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0969698924004107.

[23] W. Ali, "Optimizing the e-waste management in india: A sustainable mathematical modeling approach to circular economy," *Quality & Quantity*, pp. 1–32, 2025. DOI: https://doi.org/10.1007/s11135-025-02176-w.

[24] C. Sun, M. Xu, and B. Wang, "Deep learning: Spatiotemporal impact of digital economy on energy productivity," *Renewable and Sustainable Energy Reviews*, vol. 199, p. 114 501, 2024, ISSN: 1364-0321. DOI: https://doi.org/10.1016/j.rser.2024.114501. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032124002247.

[25] A. Gehlot, N. Sidana, D. Jawale, N. Jain, B. P. Singh, and B. Singh, "Technical analysis of crop production prediction using machine learning and deep learning algorithms," in *2022 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*, 2022, pp. 1–5. DOI: 10.1109/ICSES55317.2022.9914206.

[26] S. Mukherjee, B. Sadhukhan, N. Sarkar, D. Roy, and S. De, "Stock market prediction using deep learning algorithms," *CAAI Transactions on Intelligence Technology*, vol. 8, no. 1, pp. 82–94, 2023. DOI: https://doi.org/10.1049/cit2.12059.

[27] T. H. H. Aldhyani and A. Alzahrani, "Framework for predicting and modeling stock market prices based on deep learning algorithms," *Electronics*, vol. 11, no. 19, 2022, ISSN: 2079-9292. DOI: 10.3390/electronics11193149. [Online]. Available: https://www.mdpi.com/2079-9292/11/19/3149.

[28] 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.

[29] B. Abdollahzadeh et al., "Puma optimizer (PO): A novel metaheuristic optimization algorithm and its application in machine learning," en, *Cluster Computing*, vol. 27, no. 4, pp. 5235–5283, Jul. 2024, ISSN: 1573-7543. DOI: 10.1007/s10586-023-04221-5. Accessed: Oct. 8, 2025.

[30] S. Mirjalili, "SCA: A Sine Cosine Algorithm for solving optimization problems," *Knowledge-Based Systems*, vol. 96, pp. 120–133, Mar. 2016, ISSN: 0950-7051. DOI: 10.1016/j.knosys.2015.12.022.