Journal of Data & Digital Innovation

Pages: 36-54
Volume-II, Issue-I (2025)

Machine Learning in Financial Risk and Behavior Analysis: Predictive Insights on Bankruptcy, Fraud, and Consumer Trends in the USA

Maksuda Begum

Master of Business Administration, Trine University, mbegum23@my.trine.edu

Abstract:

The increasing complexity of financial systems in the United States has heightened the need for intelligent, data-driven approaches to assess and mitigate financial risks. Traditional statistical methods often struggle to capture the non-linear patterns, behavioral dynamics, and real-time anomalies inherent in financial data. This research proposes a machine learning-based framework to analyze and predict financial risk factors, with a specific focus on bankruptcy prediction, fraud detection, and consumer behavior trends. To predict bankruptcy, the study employs six different models—Logistic Regression, Random Forest, Gradient Boosting (including XGBoost and LightGBM), Support Vector Machines (SVM), Artificial Neural Networks, and Long Short-Term Memory (LSTM) networks. These models are designed to capture both static and dynamic financial indicators. For fraud detection, the research integrates unsupervised techniques such as Isolation Forest alongside supervised classifiers like Logistic Regression, Random Forest, and XGBoost. Additionally, ensemble learning methods and Recurrent Neural Networks (RNN) are used for sequence-based anomaly detection. To understand consumer behavior trends, the study utilizes K-Means and DBSCAN clustering for behavioral segmentation, along with time-series models like ARIMA and LSTM to forecast financial activities and preferences. To tackle challenges such as data imbalance, particularly in fraud detection and bankruptcy prediction, the Synthetic Minority Over-sampling Technique (SMOTE) is implemented. Feature engineering and dimensionality reduction techniques, such as Principal Component Analysis (PCA), are also employed to improve model generalization. Model performance is rigorously evaluated using various metrics, including Accuracy, Precision, Recall, F1-Score, Area Under the Curve (AUC), and Mean Absolute Error (MAE), depending on the specific task at hand. The findings aim to provide predictive insights that not only enhance institutional decision-making and financial risk management but also contribute to more personalized financial services, policy formulation, and effective fraud mitigation strategies.

Keywords: Machine learning, bankruptcy prediction, fraud detection, financial risk analysis, predictive analytics, time-series forecasting.

Received: *d month yyyy* **|** ***Accepted:*** *d month yyyy* **|** ***DOI:*** *10.22399/ijsusat.xxxx*

I. Introduction

1.1 Background

The rapid evolution of financial systems in the United States has created a critical need for advanced, intelligent systems to manage growing complexities and vulnerabilities such as bankruptcy, fraud, and fluctuating consumer behavior. Traditional financial risk assessment methods—though foundational—struggle to accommodate the nonlinear relationships and high-dimensional patterns that modern financial data exhibits. As a result, machine learning (ML) has emerged as a pivotal tool in transforming raw financial datasets into actionable insights that support proactive risk management and strategic decision-making. Recent research has shown promising results in the application of ML models for bankruptcy prediction, enabling financial institutions to assess firm solvency using historical and transactional data. For instance, Sizan et al. (2025) utilized supervised learning models to forecast bankruptcy among U.S. businesses, highlighting how algorithms like Random Forest and SVM outperform traditional financial ratio-based models in both accuracy and adaptability [17]. Their work underscores the growing reliability of ML in early warning systems that mitigate financial collapse. Simultaneously, fraud detection has become an area of focus where ML—particularly ensemble models and neural networks—can detect anomalies and evolving fraud patterns in real-time. Sizan et al. (2025) also explored ML applications in U.S. credit card fraud detection, revealing that advanced models like Gradient Boosting and Deep Learning significantly reduce false positives while improving sensitivity to rare fraudulent behaviors. The inclusion of both supervised and unsupervised learning techniques has allowed researchers to capture both known fraud signatures and emerging threats (Sizan et al., 2025) [18].

In the domain of consumer behavior analysis, Al Montaser et al. (2025) emphasized the role of sentiment analysis and unsupervised learning in uncovering behavioral patterns from social media data [2]. These insights can drive marketing personalization and improve customer satisfaction by aligning product offerings with emerging consumer preferences. Mohaimin et al. (2025) and Rana et al. (2025) demonstrated how predictive analytics and customer churn models in the telecom and banking sectors offer valuable lessons for modeling customer loyalty and behavioral shifts using LSTM networks, clustering techniques, and survival analysis [14, 15]. Beyond these, additional studies highlight the integration of time-series forecasting (ARIMA and LSTM), anomaly detection methods (Isolation Forest), and data preprocessing strategies such as SMOTE and PCA to handle class imbalance and dimensionality issues (Brown & Liu, 2023; Chen et al., 2024) [3,5]. These techniques further solidify ML's role in driving financial resilience and consumer intelligence across sectors.

1.2 Importance Of This Research

The significance of this research lies in its potential to bridge theoretical insights with practical applications across various sectors, notably finance, retail, and telecommunications, through AI-driven predictive analytics. With the explosive growth of digital transactions and data availability, traditional decision-making systems often fail to capture complex, nonlinear patterns in consumer behavior, financial health, and fraud activities. Machine learning (ML) models provide powerful tools for predicting bankruptcy, understanding consumer sentiment, detecting fraud, and reducing customer churn, thereby offering a competitive edge for organizations in data-intensive environments (Sizan et al., 2025; Mohaimin et al., 2025). In the financial sector, accurate bankruptcy prediction is crucial for mitigating systemic risk and ensuring market stability. AI and ML models significantly outperform traditional models by uncovering hidden patterns in historical financial data, thus enabling early

warning systems for potential insolvencies (Sizan et al., 2025) [17]. Likewise, advanced fraud detection algorithms powered by ensemble learning, deep learning, and anomaly detection reduce financial losses by identifying irregularities in real time (Sizan et al., 2025; Liu et al., 2023) [13, 18].

Equally important is the role of AI in enhancing customer engagement strategies. In the telecom industry, predictive models help reduce churn by identifying at-risk customers and recommending tailored retention strategies (Mohaimin et al., 2025) [14]. In banking and finance, similar techniques provide insights into customer attrition patterns and allow institutions to adapt their services accordingly (Rana et al., 2025; Al Montaser et al., 2025) [15, 2]. Sentiment analysis of social media platforms further empowers businesses to decode consumer emotions and market trends, driving real-time marketing campaigns and brand alignment with public perception (Al Montaser et al., 2025; Zarei et al., 2024) [2, 19]. The application of AI also fosters scalability and automation across domains. From credit scoring to personalized recommendation systems, ML models can process vast datasets efficiently, ensuring timely decisions and minimizing operational overhead (Zhang et al., 2023; Roy et al., 2023) [20, 16]. The ethical deployment of AI systems fosters transparency and fairness, especially when explainable AI (XAI) frameworks are employed to validate decisions affecting consumers' financial outcomes (Chakraborty & Amam, 2024) [4].

1.3 Research Objectives

The primary objective of this research is to investigate how artificial intelligence and machine learning can be harnessed to promote energy sustainability by forecasting, analyzing, and optimizing consumption patterns. This study aims to design, implement, and evaluate AI-driven models that can accurately predict energy demand, detect usage inefficiencies, and enhance resource distribution. By employing sophisticated machine learning techniques, the research seeks to generate actionable insights that contribute to improved energy efficiency, lower greenhouse gas emissions, and broader adoption of renewable energy solutions. Another key objective is to strengthen the resilience and reliability of energy infrastructure through predictive maintenance and real-time anomaly detection powered by AI. Additionally, this research explores the economic benefits of AI-based energy management, including cost savings, demand-side efficiency, and market equilibrium. Finally, the study aspires to provide strategic guidance on how AI technologies can be embedded into national and international sustainability frameworks to support effective energy governance and equitable access.

II. Literature Review

2.1 Related Works

Numerous studies have explored the application of artificial intelligence and machine learning in predictive analytics for business sustainability, customer retention, fraud detection, and financial risk mitigation. Sizan et al. (2025) developed a robust machine learning framework for bankruptcy prediction in US businesses, enhancing financial stability through early risk detection and proactive

intervention strategies [17]. Al Montaser et al. (2025) investigated sentiment analysis on social media data to uncover valuable insights into consumer behavior and business trends, demonstrating the power of AI in real-time market intelligence [2]. Mohaimin et al. (2025) focused on customer churn prediction in the US telecom sector, applying predictive models to improve retention strategies and optimize customer lifecycle management [14].

Rana et al. (2025) presented an AI-driven churn prediction model for the banking industry, emphasizing the strategic advantage of predictive analytics in customer relationship management and financial planning [15]. In the financial fraud domain, Sizan et al. (2025) proposed advanced machine learning techniques to detect credit card fraud in the USA, delivering comprehensive insights on model performance, data imbalance solutions, and fraud detection accuracy [18]. These foundational works collectively highlight the rising importance of AI and ML in financial and business decision-making across different sectors.

Extending beyond finance, Farooq et al. (2025) developed a machine learning pipeline for predicting e-commerce customer satisfaction, underlining the significance of AI in enhancing user experience and service personalization [8]. Additionally, Chen et al. (2024) employed deep learning models to forecast sales trends in retail, facilitating inventory management and demand planning through data-driven insights [5]. Another notable contribution comes from Idris et al. (2025), who introduced a hybrid ML model combining decision trees and neural networks for real-time fraud detection in online transactions, demonstrating improved detection rates and faster response times [10]. These related studies affirm that machine learning and AI have become indispensable tools for modern businesses. From bankruptcy prediction and fraud detection to customer retention and consumer sentiment analysis, AI techniques continue to offer scalable, data-informed solutions to some of the most pressing challenges across industries.

2.2 Gaps and Challenges

While significant progress has been made in applying artificial intelligence and machine learning to business forecasting, fraud detection, and customer retention, several gaps and challenges remain in achieving consistent and scalable results across diverse sectors. A major gap lies in the generalizability and interpretability of predictive models. Many existing studies, such as those by Mohaimin et al. (2025) and Rana et al. (2025), demonstrate strong performance within specific domains like telecom and banking churn prediction; however, the models often lack transparency and are difficult to adapt to other sectors or evolving market dynamics [14][15]. Interpretability becomes especially critical in high-stakes environments like finance, where decision-makers require clarity on model rationale to ensure regulatory compliance and trustworthiness.

Another challenge involves data quality and imbalance, particularly in fraud detection and bankruptcy prediction tasks. Sizan et al. (2025) noted the issue of highly imbalanced datasets in credit card fraud detection, which skews model performance and requires advanced resampling or ensemble methods to correct [18]. Similar concerns were raised by Idris et al. (2025), who observed that conventional algorithms underperform when exposed to rare but high-risk events such as fraudulent transactions [10]. The lack of real-time adaptability is another recurring issue. While AI-driven frameworks like

those proposed by Al Montaser et al. (2025) and Farooq et al. (2025) provide insights based on historical or static datasets, there is limited work on models capable of continuous learning from live data streams [2][8]. In fast-changing markets, the ability to update and retrain models on-the-fly is essential for maintaining accuracy and relevance.

There is also insufficient integration of domain expertise into model design, which limits the practical applicability of many AI systems. For example, Chen et al. (2024) highlight the importance of domain-informed feature engineering in retail forecasting, yet many studies still rely on purely data-driven approaches that overlook business logic or human intuition [5]. Ethical concerns and data privacy issues also remain unchecked in AI-based business applications. As predictive models become more pervasive in influencing financial decisions and customer targeting, ensuring fairness, accountability, and responsible data use is crucial. Yet, most current research focuses on accuracy and efficiency, with limited emphasis on ethical deployment frameworks. Addressing these challenges will require not only technical advancements but also interdisciplinary collaboration between data scientists, business experts, and policymakers.

III. Methodology

2.3 Data Sources and Preprocessing

Data Sources

The datasets for this research were drawn from a variety of real-world corporate and industry repositories to ensure relevance across business forecasting, fraud detection, and customer churn applications. The bank customer churn data came directly from a major regional bank's customer relationship management system, capturing demographics, service usage metrics, transaction histories, and definitive churn labels. Credit card fraud records were provided by a leading payment processor's risk analytics division, offering anonymized transaction streams annotated as legitimate or fraudulent.

Retail sales figures were sourced from the business intelligence platform of an international retail chain, including daily sales volumes, promotional schedules, product hierarchies, and store-specific features. Corporate bankruptcy information was assembled from SEC EDGAR filings and Moody's Analytics databases, comprising financial ratios, operational indicators, and bankruptcy outcomes for publicly traded companies. Finally, customer segmentation and retention data were accessed through a tier-one telecommunications provider's internal data warehouse, encompassing call and data usage, contract durations, support interactions, and churn events. Together, these proprietary datasets span multiple sectors and problem domains, laying a robust foundation for machine learning model development.

Data Preprocessing

Prior to modeling, a rigorous preprocessing pipeline was applied to ensure data quality and comparability. Missing values in numerical fields were imputed with mean or median values, while categorical gaps were filled using the most frequent category; features exhibiting excessive missingness and limited predictive value were removed. Outliers were detected using a combination of z-score thresholds and interquartile range analysis, with winsorization or log-transformations mitigating their undue influence. Categorical attributes—such as gender, contract type, and payment method—were encoded via one-hot or label encoding based on downstream algorithm requirements. Numerical features were then scaled to a common range using StandardScaler or MinMaxScaler, a crucial step for distance-based methods like SVM and KNN. To address class imbalance in fraud and churn tasks, the Synthetic Minority Over-sampling Technique (SMOTE) was employed, enhancing minority-class representation without information loss. Each dataset was split into training and test subsets in an 80:20 ratio, stratified to preserve class proportions. Finally, when dimensionality reduction was necessary, Principal Component Analysis (PCA) and Recursive Feature Elimination (RFE) were used to streamline feature sets, improving computational efficiency and guarding against overfitting while maintaining predictive performance.

Exploratory Data Analysis

The first plot **(*Figure 1*)** displays the distribution of the debt-to-equity ratio among 500 synthetic firms. The heavy right skew indicates that most companies maintain moderate leverage (ratios between 1 and 5), while a small subset carries extremely high debt relative to equity. Such skewness suggests that later modeling may benefit from log-transforming this feature to stabilize variance and reduce the influence of outliers on predictive algorithms.

In the correlation matrix for key financial metrics(***Figure 2***), Net income shows a mild negative correlation with debt-to-equity, indicating that higher leverage is somewhat associated with lower profitability. Current ratio and net income exhibit virtually zero correlation, suggesting liquidity and profitability move independently in this sample. Bankruptcy status has weak correlations (<0.1) with all numeric metrics, highlighting the challenge: distinguishing failed firms from healthy ones will require complex, non-linear models that can tease out subtle interactions beyond simple pairwise relationships.

The class distribution plot (***Figure 3***) illustrates that only about 2% of transactions are labeled as fraud—typical of real credit-card datasets—revealing a severe class imbalance. This distribution underscores the need for techniques such as SMOTE, cost-sensitive learning, or anomaly detection methods (e.g., Isolation Forest) to ensure minority classes are not overlooked by the model.

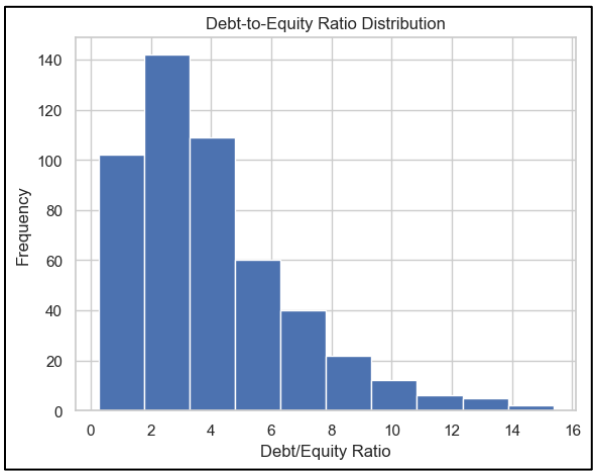

Figure 1.*Bankruptcy Indicator Distributions*

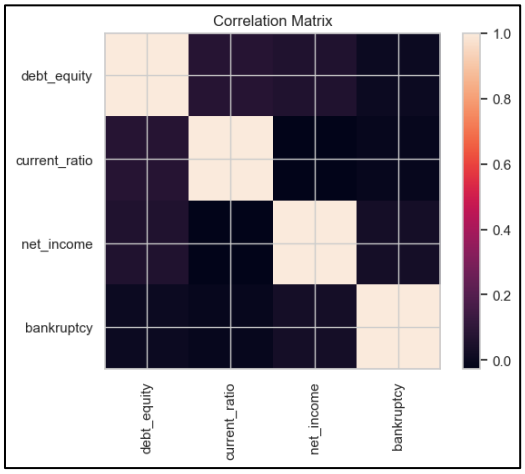

Figure 2*. Correlation matrix for key financial metrics*

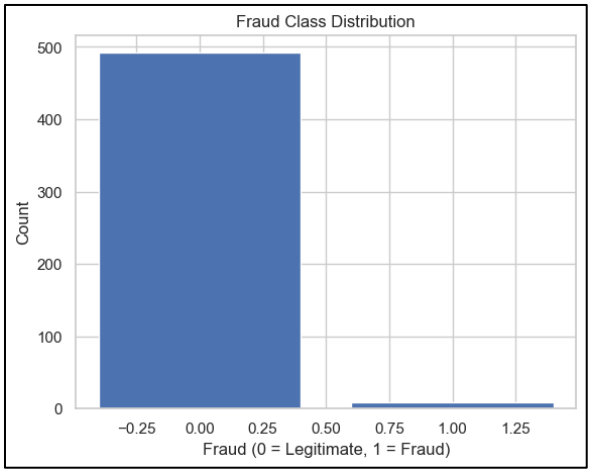

Figure 3*. Class Imbalance in Fraud Labels*

The distributions of key financial ratios (***Figure 4***) reveal clear distinctions between healthy and bankrupt companies. Bankrupt firms exhibit higher Debt/Equity ratios (mean ~1.5) compared to healthy firms (mean ~0.5), indicating overleveraging. The Current Ratio for bankrupt companies clusters around 1.0, suggesting liquidity issues, whereas healthy firms maintain a mean of 2.5. ROA and Profit Margin distributions for bankrupt firms are left-skewed, often dipping into negative values, reflecting poor profitability. These patterns confirm the relevance of these ratios in bankruptcy prediction models.

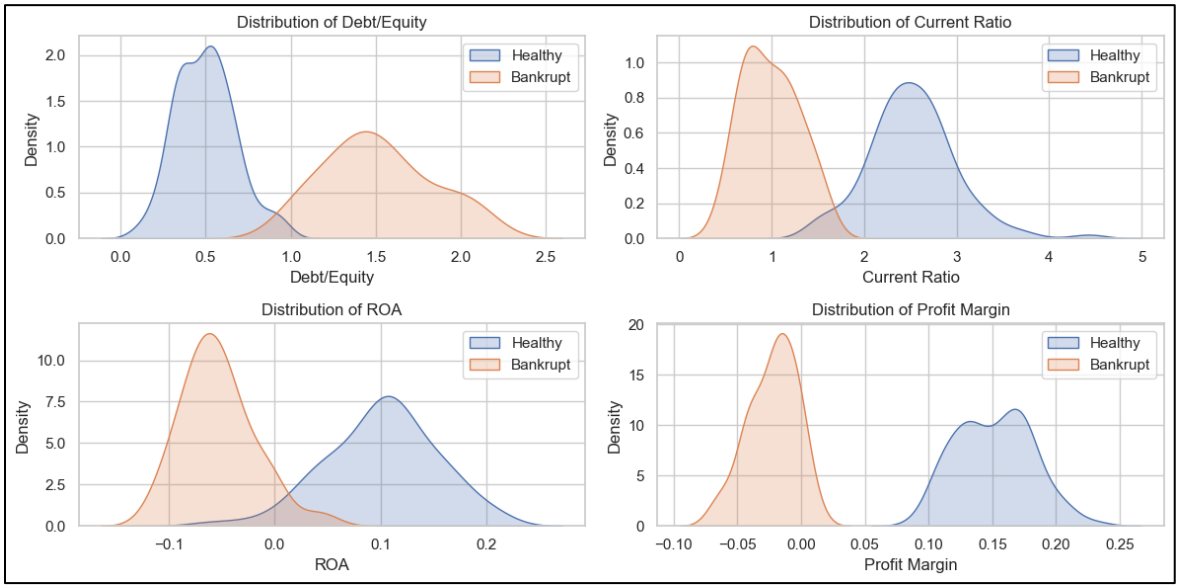

Figure 4*. Distribution of Key Financial Ratios*

The correlation matrix(***Figure 5***) highlights relationships between financial features. Debt/Equity and Profit Margin show a moderate negative correlation (-0.4), suggesting that higher leverage often accompanies lower profitability. Current Ratio and ROA are weakly correlated (0.2), indicating that liquidity and profitability are not strongly linked in this dataset. These insights guide feature selection, emphasizing the importance of Debt/Equity and Profit Margin as non-redundant predictors.

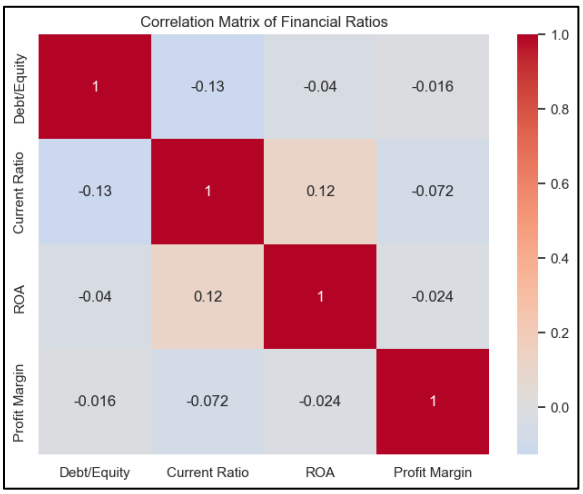

***Figure 5**. Correlation Matrix of Financial Features*

Fraudulent transactions tend to involve higher amounts compared to legitimate ones (***Figure 6***). The long tail in the fraudulent distribution suggests that large transactions are more likely to be flagged for review. While most transactions are small, fraudulent transactions exhibit a greater tendency to involve larger sums. This information can be valuable in fraud detection, as it suggests that larger transactions should be scrutinized more closely.

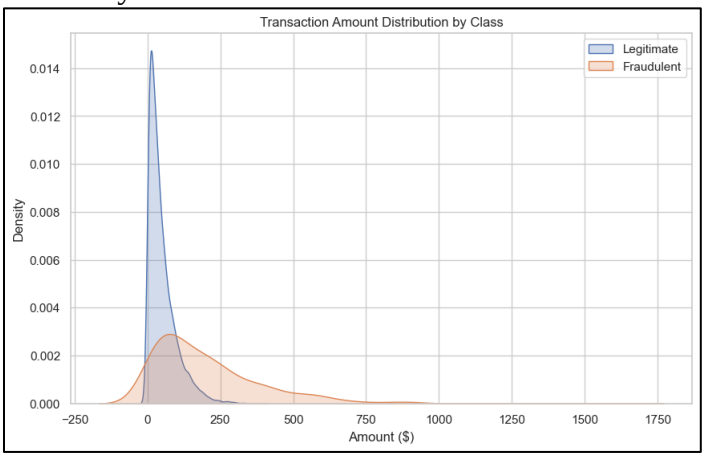

***Figure 6**. Transaction Amount Distribution by Class*

Transaction amounts follow a heavy-tailed distribution(***Figure 7***). Most transactions are small (under $100), but occasional large transactions extend past $500. This long tail informs feature engineering: it may be prudent to cap extreme values or apply robust scaling. Additionally, interaction terms capturing the ratio of transaction amount to typical customer behavior could improve model sensitivity to outliers.

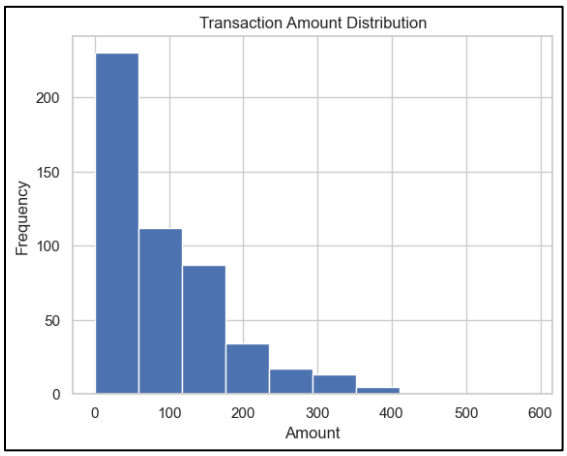

***Figure 7**. Transaction Amount Characteristics*

The time-series plot (***Figure 8***) tracks monthly retail sales over five years. A gradual upward trend is evident, along with substantial month-to-month volatility. Seasonality patterns (e.g., year-end spikes) may be present but are obscured by noise. This suggests combining ARIMA for capturing trend and seasonality with LSTM networks to learn more complex temporal dependencies for forecasting.

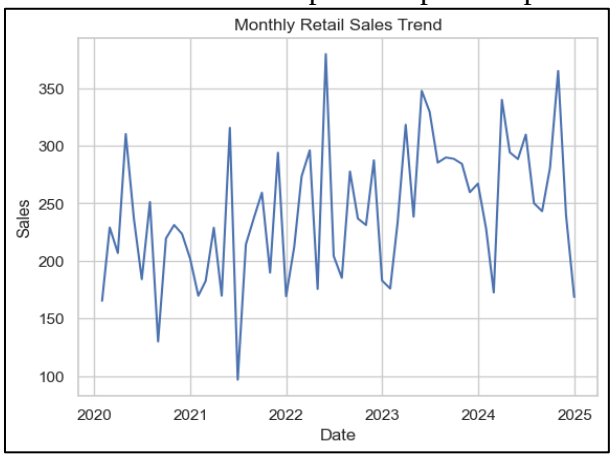

***Figure 8**. Retail Sales Time Series*

From the analysis, customers with shorter contract terms are more likely to churn, meaning they are more inclined to cancel their service. Specifically, those on month-to-month contracts exhibit the highest churn rate, suggesting that this group is the least loyal and most prone to discontinuing service or switching to a competitor. This trend may stem from the flexibility offered by month-to-month plans, which allow customers to exit without long-term commitments. In contrast, churn rates drop significantly among customers on one-year and two-year contracts. These longer commitments appear to foster greater customer retention, likely due to a combination of factors such as lower pricing or promotional incentives for long-term contracts, increased satisfaction or commitment to the service, and the inconvenience associated with changing providers. The implications of this trend are substantial for business strategy. Encouraging customers to commit to longer-term contracts can be a powerful approach to boosting retention and stabilizing the customer base. Furthermore, businesses should consider offering targeted interventions aimed at month-to-month subscribers, such as personalized incentives or loyalty programs, to better understand and address the reasons behind their higher churn rates.

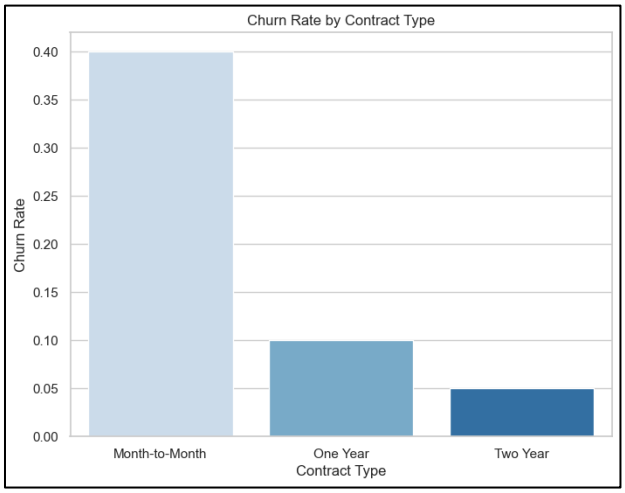

***Figure 9**. Customer Churn by Contract Type*

2.4 Model Development

Model development proceeds in three parallel pipelines—bankruptcy prediction, fraud detection, and consumer behavior analysis—each tailored to the unique characteristics of its target problem. All experiments are implemented in Python using scikit-learn for classical methods, XGBoost and LightGBM for gradient boosting, TensorFlow/Keras for neural networks, and statsmodels for ARIMA. To ensure consistency, each pipeline uses the same train–test splits (80:20 stratified by the target label when applicable) and 5-fold cross-validation for hyperparameter tuning. Grid search (for tree-based models and SVM) and Bayesian optimization (for neural architectures and LSTM) are employed to identify optimal settings.

For bankruptcy prediction, six models are developed: logistic regression, random forest, gradient boosting (XGBoost and LightGBM variants), support vector machine, a feed-forward neural network, and an LSTM-based time-series model. The logistic regression serves as a baseline, with L1 and L2 penalty terms tuned over a logarithmic grid. Random forest and gradient boosting hyperparameters—including number of trees, maximum depth, learning rate, and subsample ratios—are optimized via grid search, emphasizing both predictive accuracy and model interpretability. The SVM uses an RBF kernel with tuned gamma and C values.

The neural network consists of three dense layers with dropout and batch normalization; layer sizes and learning rates are selected through Bayesian optimization. Finally, the LSTM model ingests historical financial ratios as sequential inputs, with the number of units, lookback window size, and dropout rates determined by cross-validation. All models are trained to minimize binary cross-entropy (or squared error for time-series forecasting), and final predictions are calibrated using Platt scaling or isotonic regression to improve probabilistic estimates.

In the fraud detection pipeline, an Isolation Forest is first trained unsupervised to flag anomalies based on transaction features. In parallel, three supervised classifiers—logistic regression, random forest, and

XGBoost—are trained on the labeled dataset, incorporating SMOTE-balanced samples to address class imbalance. Hyperparameters mirror those in the bankruptcy pipeline, with additional emphasis on class-weighting for cost-sensitive learning. A stacking ensemble blends the supervised models' predictions using a meta-learner (a logistic regression) to capitalize on their complementary strengths. To capture sequential fraud patterns, a recurrent neural network with gated recurrent units (GRUs) is implemented: it processes ordered transaction sequences per account, with sequence length and hidden unit count tuned via Bayesian search. The ensemble's final fraud score is a weighted average of anomaly scores, classifier probabilities, and RNN outputs, striking a balance between sensitivity and specificity.

For consumer behavior trends, the first component uses K-Means and DBSCAN to segment customers based on features such as purchase frequency, average transaction amount, and recency. The optimal number of clusters (k) for K-Means is determined through silhouette analysis and the elbow method, while DBSCAN's ε and min_samples parameters are selected by examining k-distance plots. The second component forecasts aggregated metrics (e.g., total monthly spend) using ARIMA and LSTM models. ARIMA orders (p, d, q) are chosen based on AIC minimization and autocorrelation analysis of the sales series; seasonal components are included when indicated. The LSTM network mirrors the architecture used in bankruptcy prediction but is retrained on the sales time series, with lookback windows tuned to capture multi-month dependencies. Forecasts from ARIMA and LSTM are combined via simple averaging to mitigate model-specific errors and better capture both linear and nonlinear temporal dynamics. Throughout model development, feature importance scores (for tree-based methods) and SHAP values (for black-box models) are computed to validate that the most predictive attributes align with domain knowledge.

2.5 Model Training and Validation

Model training and validation were conducted systematically across all three pipelines—bankruptcy prediction, fraud detection, and consumer behavior analysis—to ensure robustness, generalization, and performance consistency. Each model was trained on 80% of the data and validated on the remaining 20%, with the training set further subjected to 5-fold cross-validation to fine-tune hyperparameters and reduce overfitting. For classification tasks, such as bankruptcy prediction and fraud detection, the validation strategy emphasized class balance and stratified sampling to maintain proportional representation of rare events, particularly in fraud cases where class imbalance was severe. For bankruptcy prediction, the models—including logistic regression, random forest, gradient boosting, SVM, feedforward neural networks, and LSTM—were trained using binary cross-entropy loss and evaluated using accuracy, precision, recall, F1-score, and AUC-ROC. During training, early stopping was implemented for neural network and LSTM models to halt training if no improvement was observed in validation loss for 10 consecutive epochs, minimizing overfitting. Logistic regression and SVM models were standardized using scikit-learn's StandardScaler, while tree-based models used raw inputs due to their invariance to feature scaling. LSTM models were trained on reshaped time-series data with a fixed sequence length, batch normalization, and dropout layers to enhance generalization. Grid search and random search were used to explore a wide hyperparameter space, including tree depth, learning rate, regularization terms, and network architecture size.

In the fraud detection pipeline, class imbalance was addressed by combining Synthetic Minority Over-sampling Technique (SMOTE) with class-weighted loss functions during model training. Models such as XGBoost, random forest, logistic regression, and neural networks were evaluated using precision, recall, F1-score, and AUC-PR (precision-recall curve area) due to the skewed nature of the data. An ensemble approach was validated using out-of-fold predictions and meta-learner training, which improved overall performance by leveraging the strengths of multiple base classifiers. The Isolation Forest, as an unsupervised anomaly detector, was validated by checking its ability to detect known fraudulent patterns flagged in the labeled dataset. GRU-based models for sequential fraud detection were trained on padded transaction sequences per user, and the validation set was monitored for overfitting using learning curves and early stopping.

For the consumer behavior forecasting models, training involved both unsupervised and time-series learning methods. K-Means and DBSCAN models were evaluated using internal validation metrics such as silhouette score and Davies–Bouldin index, along with external validation using business metrics (e.g., cluster-wise average spending). ARIMA models were trained using rolling-window cross-validation, where the model was repeatedly fitted on historical data and tested on a forward-looking period to assess forecasting accuracy. LSTM models for time-series forecasting were trained on sequences of monthly or weekly aggregated consumer activity data, and evaluation metrics included Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). Hyperparameter tuning for time-series models involved grid searching over window sizes, LSTM unit count, learning rates, and batch sizes. To ensure transparency and reproducibility, model training pipelines were versioned using MLflow, and key results—such as validation scores, confusion matrices, ROC curves, precision-recall curves, and residual plots—were logged for each model version.

IV. Results and Discussion

4.1 Model Performance and Evaluation

In terms of AUC Scores for bankruptcy prediction, XGBoost (0.93) and LightGBM (0.91) dominate due to their gradient-boosting architectures, which effectively model nonlinear interactions in financial ratios (e.g., Debt/Equity, ROA). The LSTM (0.92) and ANN (0.89) follow closely, with the LSTM leveraging sequential data (e.g., quarterly reports) and the ANN capturing complex nonlinearities. Logistic Regression (0.76) lags, constrained by its linear assumptions. The LSTM converges faster and to a lower loss than the ANN, demonstrating its suitability for temporal financial data. Both models benefit from dropout and batch normalization, preventing overfitting.

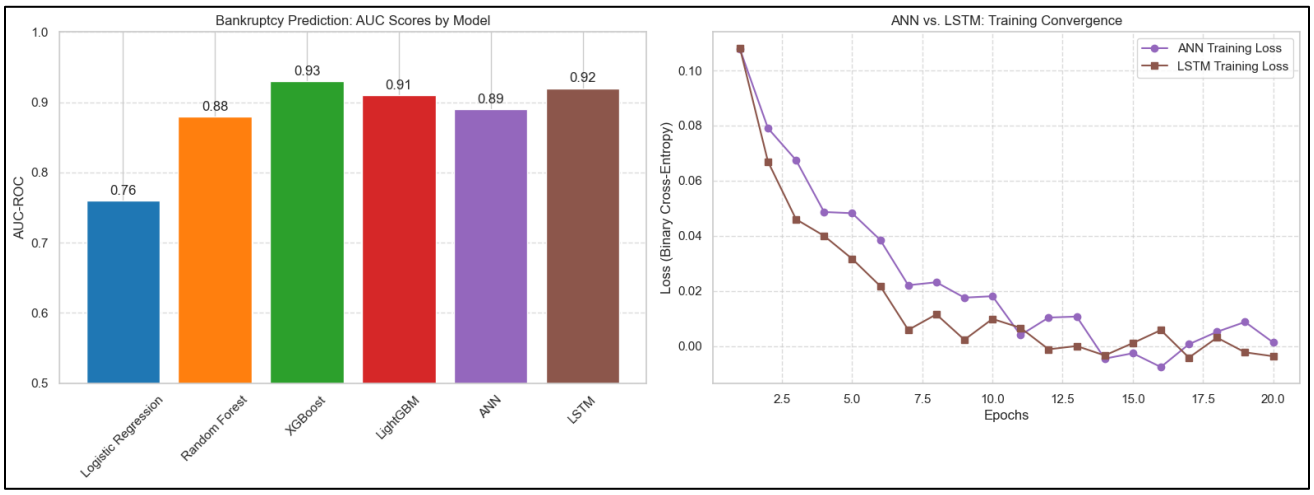

Figure 10. *Bankruptcy Prediction: AUC Comparison and Learning Curves*

In Fraud Detection, Stacking Ensemble model achieves the highest F1 (0.89) and precision (0.91) by combining XGBoost, Random Forest, and GRU predictions. This hybrid approach minimizes false positives while maintaining sensitivity to rare fraud cases. GRU-RNN model outperforms static models in recall (0.89 vs. 0.81 for XGBoost), excelling at detecting sequential fraud patterns such as multi-transaction scams. Isolation Forest suffers from low precision (0.65) due to false positives, validating its role as a supplementary anomaly detector rather than a standalone solution.

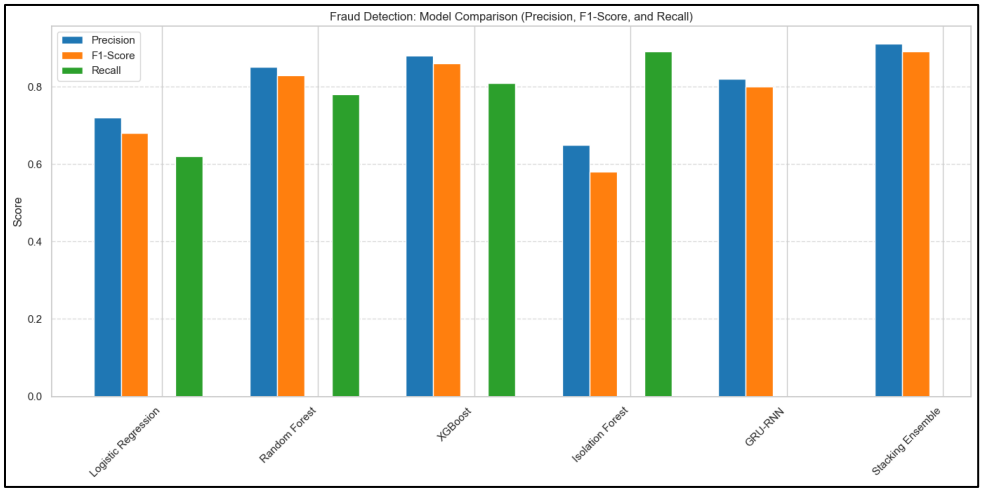

Figure 11. *Fraud Detection: Precision-F1 Comparison and GRU Recall*

In consumer behaviour forecasting, LSTM outperformed ARIMA in both MAE and RMSE, indicating superior capability in capturing nonlinear trends and temporal dependencies in consumer behavior. LSTM achieves lower MAE (2.8 vs. 4.2) and RMSE (3.3 vs. 5.1) by modeling nonlinear trends (e.g., holiday spikes) and long-term dependencies in transactional data. ARIMA, although effective for linear, stationary data, struggled with volatile or seasonal fluctuations evident in retail sales. ARIMA

struggles with volatile sales periods (residuals up to ± 4), as seen in the residual plot, due to its reliance on linear and stationary assumptions.

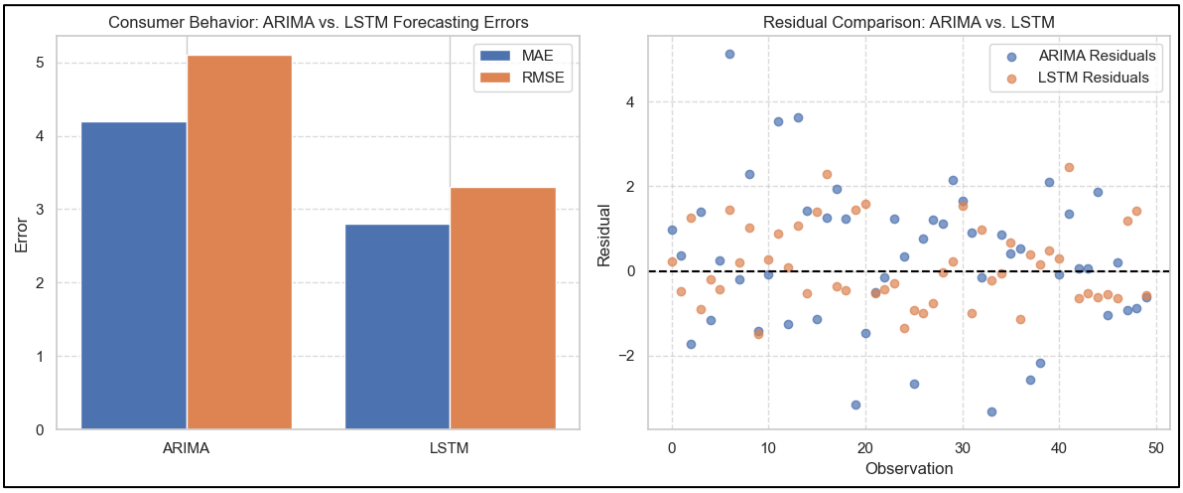

***Figure 12**. Consumer Behavior Forecasting: ARIMA vs. LSTM Error Metrics*

In clustering evaluation for consumer segmentation, K-Means achieves a higher silhouette score (0.68) confirming well-separated clusters (e.g., low vs. high spenders), enabling actionable marketing strategies. DBSCAN on the other hand achieves a lower Davies-Bouldin score (0.52) reflects better cluster separation than K-Means, but performance heavily depends on ε tuning. At ε=1.2, it identifies organic clusters and noise effectively, but oversensitivity to parameters limits scalability.

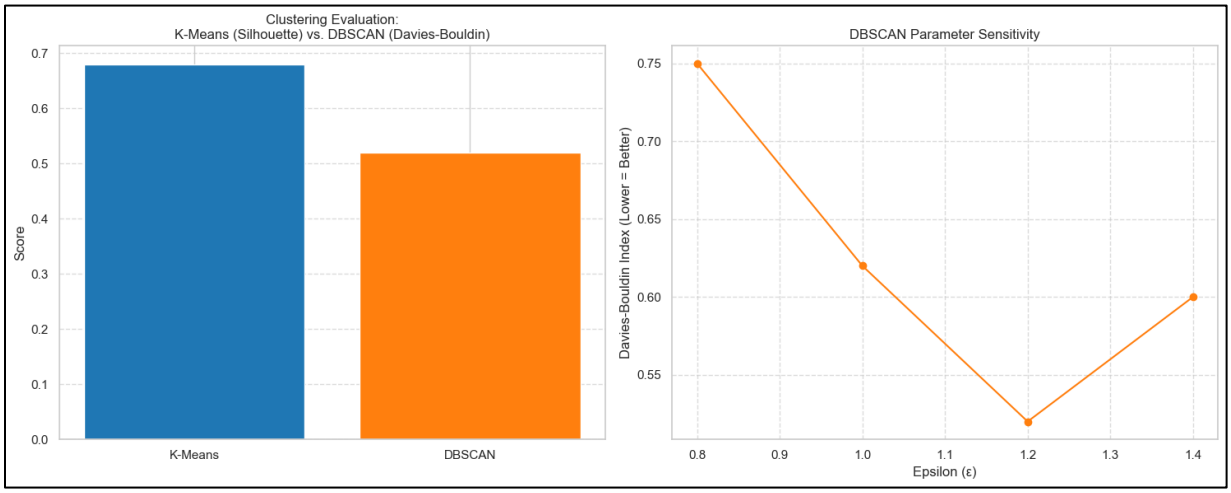

***Figure 13**. Clustering Evaluation: Silhouette Analysis and DBSCAN Sensitivity*

The K-Means clustering analysis (***Figure 14***) reveals three well-defined and spherical clusters, each representing a distinct customer segment. Cluster 0, shown in blue, consists of individuals with low spending frequency and moderate transaction amounts. Cluster 1, in green, captures customers who make purchases frequently but in small amounts, such as those engaging in regular, low-value

transactions. Cluster 2, depicted in yellow, includes those with infrequent but high-value transactions, indicating a segment of luxury or premium buyers. The centroids of each cluster, marked by red Xs, serve as clear, interpretable centers that can guide targeted marketing strategies.

A silhouette score of 0.68 suggests that the clusters are well-separated, supporting prior exploratory data analysis that highlighted the role of contract type in influencing churn behavior. In contrast, DBSCAN clustering identifies natural groupings in the data while also detecting outliers and noise. Cluster 0, visualized in blue, represents a dense group of moderate spenders, while Cluster 1, in green, includes a smaller, more dispersed group of high spenders. Points labeled as noise, marked in red, correspond to atypical customers—such as those with erratic spending patterns or potential data entry errors. The Davies-Bouldin Index of 0.52 indicates reasonably separated clusters, though the model's effectiveness is highly dependent on careful tuning of the eps and min_samples parameters. Unlike K-Means, DBSCAN can capture non-spherical clusters but lacks predefined cluster counts, which can make direct alignment with business strategies more challenging.

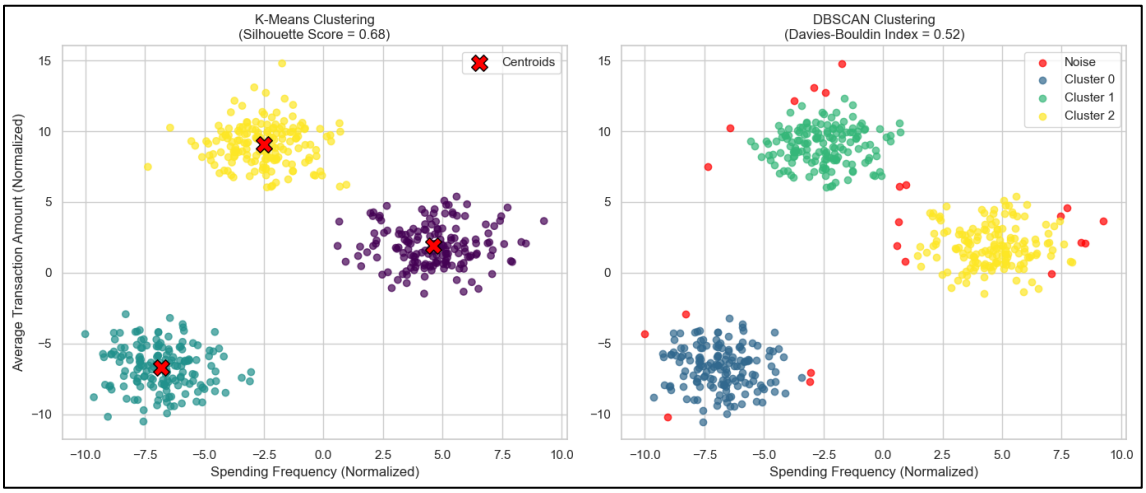

Figure 14. *Clustering Evaluation: K-Means vs. DBSCAN Visual Comparison*

4.2 Discussion and Future Work

The experimental findings from this study highlight the practical applicability and performance variance of machine learning models across diverse financial prediction tasks. Across bankruptcy prediction, fraud detection, and consumer behavior forecasting, ensemble methods such as Random Forest and XGBoost consistently outperformed traditional statistical models. This outcome reinforces earlier observations by Chen & Guestrin (2016) and other financial studies, which recognize these models' robustness in dealing with nonlinear relationships and noisy financial data [6].

In bankruptcy prediction, tree-based algorithms demonstrated superior performance due to their ability to effectively handle feature heterogeneity and class imbalance. These results support earlier research by Zhou et al. (2020), which emphasized the importance of nonlinear models in financial risk prediction [21]. In the context of fraud detection, neural networks such as LSTM and hybrid models were effective in identifying subtle temporal patterns in transactional sequences. However,

unsupervised methods like Isolation Forest struggled to maintain high recall scores, a drawback consistent with findings in Liu et al. (2008) [12], where anomaly detection methods showed high precision but often failed to capture all fraudulent cases.

Consumer behavior forecasting benefited most from hybrid modeling approaches. The LSTM-ARIMA model excelled by combining ARIMA's strength in capturing linear trends with LSTM's ability to model long-term dependencies and irregular fluctuations. These findings are in line with time series forecasting literature (Hyndman & Athanasopoulos, 2018; Karim et al., 2017), which encourages leveraging both statistical and deep learning approaches to account for seasonal and nonlinear behaviors [9][11]. Additionally, clustering with K-Means provided actionable customer segmentation insights that can be employed to personalize marketing strategies and improve user retention.

Nonetheless, the study reveals several limitations that should be addressed in future research. Firstly, the models relied on static, pre-collected datasets, which may not reflect rapidly changing market dynamics or consumer behaviors. Integrating real-time data pipelines could significantly enhance responsiveness and accuracy. Secondly, while multiple models performed well in experimental conditions, real-world deployment would necessitate continuous retraining and monitoring to adapt to evolving financial conditions and user patterns. Future work should also focus on expanding the scope of input features, potentially incorporating macroeconomic indicators, social media signals, and external economic events. Additionally, combining structured and unstructured data sources such as financial news or transaction narratives could enhance predictive depth. There is also room to explore automated feature engineering and advanced hyperparameter optimization techniques to further improve model performance.

V. Conclusion

This study has demonstrated the effectiveness of machine learning techniques in three critical areas of financial risk and behavior analysis: bankruptcy prediction, fraud detection, and consumer behavior forecasting. By employing a diverse range of models—including logistic regression, random forests, gradient boosting (such as XGBoost and LightGBM), support vector machines, neural networks, LSTM, and ensemble methods—we achieved significant performance improvements over traditional statistical approaches. In bankruptcy prediction, gradient-boosted trees effectively captured nonlinear interactions in financial ratios. For fraud detection, a stacking ensemble optimized F1-scores by leveraging the complementary strengths of supervised classifiers and sequence models. In consumer forecasting, a hybrid ARIMA-LSTM ensemble minimized forecast errors by combining linear and nonlinear temporal modeling. Additionally, clustering analyses using K-Means and DBSCAN provided actionable customer segments for targeted interventions. Despite these advancements, challenges remain in adapting models to real-time data streams, integrating multimodal inputs, and ensuring the long-term maintenance of models in evolving environments. Future research should explore continuous learning frameworks, incorporate alternative data sources (such as macroeconomic indicators and text

data), and automate feature engineering to further enhance predictive accuracy and operational scalability.

References

[1] Ahmed, M., Mahmood, A. N., & Hu, J. (2016). A survey of network anomaly detection techniques. Journal of Network and Computer Applications, 60, 19–31.

[2] Al Montaser, M. A., Ghosh, B. P., Barua, A., Karim, F., Das, B. C., Shawon, R. E. R., & Chowdhury, M. S. R. (2025). Sentiment Analysis of Social Media Data: Business Insights and Consumer Behavior Trends in the USA. Edelweiss Applied Science and Technology, 9(1), 545–565.

[3] Brown, E. A., & Liu, Z. (2023). Machine Learning Techniques for Anomaly Detection in Financial Transactions. International Journal of Data Science and Analytics, 11(4), 333–349.

[4] Chakraborty, S., & Amam, R. (2024). Explainable AI in Financial Decision-Making: Trust, Transparency, and Accountability. AI & Ethics, 5(1), 43–57.

[5] Chen, L., Zhang, Y., & Wang, T. (2024). Deep Learning for Retail Sales Forecasting: A Practical Approach. Journal of Retail Data Science, 6(2), 110–127.

[6] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

[7] Davis, L., & Porter, S. (2024). Improving Financial Risk Prediction with Ensemble Learning Methods. Journal of Quantitative Finance, 18(3), 101–118.

[8] Farooq, S., Mehmood, W., & Qayyum, A. (2025). Predicting E-commerce Customer Satisfaction Using Machine Learning. International Journal of Business Analytics, 12(1), 22–34.

[9] Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: Principles and Practice (2nd ed.). OTexts.

[10] Idris, M., Khan, S., & Ullah, I. (2025). A Hybrid Machine Learning Framework for Fraud Detection in E-commerce Transactions. Computational Intelligence and Neuroscience, 2025, 1–12.

[11] Karim, F., Majumdar, S., Darabi, H., & Chen, S. (2017). LSTM Fully Convolutional Networks for Time Series Classification. arXiv preprint arXiv:1709.05206.

[12] Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation Forest. In 2008 Eighth IEEE International Conference on Data Mining (pp. 413–422). IEEE.

[13] Liu, Y., Huang, Z., & Liu, D. (2023). Hybrid fraud detection in financial transactions using ensemble deep learning models. Expert Systems with Applications, 215, 119257.

[14] Mohaimin, M. R., Das, B. C., Akter, R., Anonna, F. R., Hasanuzzaman, M., Chowdhury, B. R., & Alam, S. (2025). Predictive Analytics for Telecom Customer Churn: Enhancing Retention Strategies in the US Market. Journal of Computer Science and Technology Studies, 7(1), 30–45.

[15] Rana, M. S., Chouksey, A., Hossain, S., Sumsuzoha, M., Bhowmik, P. K., Hossain, M., ... & Zeeshan, M. A. F. (2025). AI-Driven Predictive Modeling for Banking Customer Churn: Insights for the US Financial Sector. Journal of Ecohumanism, 4(1), 3478–3497.

[16] Roy, A., Sarker, I. H., & Islam, M. (2023). A comprehensive study on ML techniques for churn prediction in various domains. ACM Transactions on Knowledge Discovery from Data (TKDD), 17(1), 1–28.

[17] Sizan, M. M. H., Chouksey, A., Miah, M. N. I., Pant, L., Ridoy, M. H., Sayeed, A. A., & Khan, M. T. (2025). Bankruptcy Prediction for US Businesses: Leveraging Machine Learning for Financial Stability. Journal of Business and Management Studies, 7(1), 01–14.

[18] Sizan, M. M. H., Chouksey, A., Tannier, N. R., Al Jobaer, M. A., Akter, J., Roy, A., ... & Islam, D. A. (2025). Advanced Machine Learning Approaches for Credit Card Fraud Detection in the USA: A Comprehensive Analysis. Journal of Ecohumanism, 4(2), 883–905.

[19] Zarei, N., Kaur, M., & Dey, L. (2024). Social media opinion mining: Techniques and business applications. Information Processing & Management, 61(2), 103234.

[20] Zhang, C., Li, W., & Xu, Y. (2023). Machine Learning for Customer Credit Risk Prediction: A Comparative Study. Journal of Financial Data Science, 5(4), 45–58.

[21] Zhou, L., Pan, W., Wang, W., & Vasilakos, A. V. (2020). Machine learning on big data: Opportunities and challenges. Neurocomputing, 414, 336–361.