TECHNICAL SCIENCES AND TECHNOLOGIES

*DOI: 10.25140/2411-5363-2025-1(39)-184-195*
*UDC 004.8:336.76:519.2*

***Mykola Zlobin[1], Volodymyr Bazylevych[2]***

[1]PhD student
Chernihiv Polytechnic National University (Chernihiv, Ukraine)
**E-mail:** mykolay.zlobin@gmail.com. **ORCID:** https://orcid.org/0009-0000-7653-6109. **Scopus Author ID:** 59337918100

[2]PhD in Economics, Associate Professor, Head of ESI EIT
Chernihiv Polytechnic National University (Chernihiv, Ukraine)
**E-mail:** bazvlamar@stu.cn.ua. **ORCID:** https://orcid.org/0000-0001-8935-446X. **Scopus Author ID:** 57214432127

# SYSTEMATIC REVIEW OF DEEP AND MACHINE LEARNING FOR FINANCIAL MODELING

*Machine learning (ML) and deep learning (DL) have impacted financial analytics, with advanced solutions for classification and regression tasks. This systematic review provides an analysis of state-of-the-art ML/DL applications in finance, with focus on methods, but also challenges. A total of 41 papers were analysed to identify trends, methodologies, and research gaps in this domain. The study begins with an overview of ML/DL in finance, main classification and regression problems, and challenges in financial data modelling. It then explores ML/DL techniques for classification tasks such as credit scoring, fraud detection, and algorithmic trading, evaluating traditional and modern approaches, including transformer-based models for sentiment analysis. Regression-oriented applications are analysed, with focus on stock price prediction, volatility forecasting, and portfolio optimization, with insights into hybrid modelling strategies. Comparative analysis assesses ML/DL models based on performance metrics, interpretability, and trade-offs between accuracy, computational complexity, and generalizability. The paper also identifies challenges, including data quality, ethical concerns, models, and the integration of ML/DL with traditional financial frameworks. Trends such as explainable AI, federated learning, and quantum computing are discussed as future directions. Findings show the increasing role of ML/DL in financial decision-making.*

***Key words:*** *Deep learning; machine learning; classification problems; regression models; financial analytics.*
*Table: 2. References: 41.*

**Urgency of the research.** Financial institutions are now dealing with datasets that traditional analytical methods struggle to process effectively. ML and DL models analyse these large volumes of data, but with improved predictive accuracy and more informed decision-making processes. These technologies have found additional applications in different financial domains, including credit scoring, fraud detection, and market forecasting [1; 2]. Practical significance of this paper represents the impact of ML and DL on financial analytics. In credit scoring, DL models have improved assessment of borrowers' creditworthiness by identifying complex patterns in financial behaviours, thereby reducing default rates. In fraud detection, ML algorithms have improved identification of fraudulent activities by analysing transaction patterns and anomalies. In market forecasting, DL techniques have improved accuracy of predicting market trends, that investors have more informed decisions [3]. The urgency of this research is highlighted by the findings in AI technologies and their potential to revolutionize financial services. Trends such as explainable AI, federated learning, and quantum computing are addressing current challenges in financial modelling, including data privacy, model interpretability, and computational limitations. Explainable AI aims to make ML models more transparent, in order to understand the decision-making process in regulated industries like finance. Federated learning enables collaborative model training across decentralized data sources, preserving data privacy, that is concern in financial data management. Quantum computing is used for solving complex optimization problems in finance at high speed, transforming risk assessment and portfolio optimization.

**Target setting.** The research addresses the challenge of systematically reviewing and synthesizing state-of-the-art ML and DL methods applied to classification and regression problems within the financial sector. Financial institutions are increasingly adopting ML and DL techniques to improve decision-making processes, risk assessment, and predictive analytics. However, the rapid evolution of these technologies presents a challenge to stay informed about the most effective and current methodologies. This research aims to systemize existing knowledge, evaluate the efficacy of different ML and DL models, and identify gaps in the current literature to guide future research directions.

**Actual scientific research and issues analysis.** Several studies have attempted to address the challenges associated with classification and regression tasks. A literature review analysed the new findings of ML and DL in financial applications, highlighting their effectiveness in different financial tasks [4]. The paper analyzed the use of DL in economics, with insights into frequently used models and their applications. This research offered a review of DL implementations in economic contexts, with both the potential and limitations of these technologies. In business and finance, a review analyzed ML applications, with the focus on techniques used in marketing, stock analysis, and demand forecasting. The paper also addressed limitations related to data quality and model interpretability [5]. Furthermore, a literature review focused on financial time series forecasting with DL, categorizing studies based on their forecasting implementation areas and model choices. This review highlighted the growing interest in DL models for financial forecasting and identified potential research opportunities [6]. Despite these analyses, challenges exist in the application of ML and DL in finance. Issues such as data imbalance, model transparency, and the need for evaluation metrics remains. Addressing these challenges is needed for the practical utility of ML and DL in financial contexts [7].

**Uninvestigated parts of general matters defining.** Despite findings in applying ML and DL to financial classification and regression tasks, some areas remain underexplored. One of the main challenges is model interpretability. Many ML and DL models operate as "black boxes", difficult to understand their decision-making processes. This lack of transparency reduces trust and slows adoption in financial institutions, where interpretability is used for regulatory compliance and risk management. Another unresolved issue is data quality. Financial datasets often contain noise, missing values, or imbalanced classes. This can introduce biases in model training. But also access to high-quality, proprietary financial data is frequently restricted. The problem of concept drift is still a major limitation in financial modelling. Market conditions are dynamic, and statistical properties change over time. Models trained on historical data may become outdated, that lead to reduced predictive accuracy.

**The research objective.** The objective of this research is to conduct a systematic review of state-of-the-art ML and DL methods applied to classification and regression problems in the financial sector. With analysing existing literature, the paper aims to identify current trends, evaluate the effectiveness of models, and uncover gaps in the application within finance. The outcome of this research is a synthesis of existing studies that analyse the strengths and limitations of ML and DL approaches in financial applications. This includes insights into model performance, data requirements, and contextual applicability. The study also aims to provide an assessment of challenges such as data quality, model interpretability. The paper serves as a resource that highlights existing research gaps, thereby guiding future investigations into ML and DL applications in finance.

**Overview of ML and DL in finance.** At first, financial institutions relied on traditional statistical methods for tasks such as risk assessment and forecasting. However, ML introduced more sophisticated algorithms capable of learning from data, that led to improved accuracy in financial modelling. As an example, decision trees and support vector machines became popular for credit scoring and fraud detection [8; 9]. The progression from traditional ML to DL marked a shift in financial analytics. DL models, specifically neural networks, have the capacity to process amounts of unstructured data, such as text and images, used for complex financial tasks. Applications of DL in finance include sentiment analysis of news articles for stock price prediction and the development of risk assessment models [10-12]. In recent years, the finance has adopted DL techniques, because of the increasing availability of big data and computational power. This trend has led to the development of more accurate predictive models and automated decision-making systems.

ML and DL have been used for addressing classification and regression problems in the financial sector. The main classification problems that have been identified are credit scoring, fraud detection, customer segmentation. ML algorithms, such as logistic regression are used in assessments of credit reliability of loan applicants. The analysis of variables like credit history, income, and employment status, these models predict the likelihood of default and helping financial institutions in making informed decisions. A study compared ML models, including neural networks and logistic regression, to predict the default status of credit card customers. The results indicated that certain models achieved high accuracy in credit risk assessment [13]. The paper [14] presents the analysis of how ML is transforming digital credit scoring, specifically in rural finance. It highlights that despite findings in financial sector, a portion of the global rural population remains underserved by traditional banking systems. According to the paper nearly one-third of the world's adult population lacks access to banking services, forcing them to rely on informal or semi-formal credit sources. The paper emphasizes how ML algorithms, including decision trees, support vector machines (SVMs), and deep learning models, are being used by financial institutions to develop alternative credit scoring models, overcoming the limitations of conventional scoring methods based on historical banking transactions. The paper also analysis the challenges of the integration of ML models with traditional banking systems, specifically regarding data availability, ethical concerns, and regulatory compliance. The paper [15] examines the accuracy trade-off in ML for credit scoring. In 2020, the total outstanding retail credit in the U.S. surpassed $4,161 billion, showing the economic importance of fair and accurate credit scoring models. The paper highlights that traditional ML models tend to optimize for profit without considering fairness, that lead to disparities in loan approvals in different demographic groups. It provides a review of fairness criteria, including independence, separation, and sufficiency, and assesses their suitability for credit scoring. Using 7 credit scoring datasets, the paper compares improved fairness approaches such as pre-processing, in-processing, and post-processing methods. The findings reveal that fair in-processing techniques have a balance between reducing bias and maintaining profitability. The paper also shows that reducing discrimination in credit scoring models can be achieved with minimal profit loss, approximately 4.91% when reducing fairness violations below a 0.2 separation metric threshold. However, achieving perfect fairness would result in over a 35% drop in profitability that makes it financially unfeasible for lenders.

DL have further improved credit scoring by identifying complex, non-linear relationships in data. An example is the development of a DL model that outperformed standard credit scoring models, even when using the same data. This model provided more accurate predictions aimed at reducing consumer default [16]. The paper [17] presents evaluation of DL in credit scoring, highlighting their advantages over traditional ML models. The paper analyses the comparison of the performance of DL architectures such as deep belief networks (DBNs), convolutional neural networks (CNNs), autoencoders (AEs), and long short-term memory (LSTM) networks. The review presents the DL applications using commonly used credit scoring datasets, including the Australian, German (categorical and numerical), Japanese, and Taiwanese datasets, assessing model accuracy through metrics like area under the receiver operating characteristic curve (AUC). The findings show that DBNs achieve higher accuracy than shallower networks, as an alternative to traditional ML models for credit scoring. Financial institutions are legally required, under regulations like the general data protection regulation and the Basel accord, to explain automated credit decisions to applicants. The paper [18] proposes a new framework that converts tabular credit scoring datasets into images, in order to enable the use of 2D CNNs for classification while ensuring model interpretability. The research applies CNN models on widely used credit datasets (German, Australian, and home equity line of credit datasets) and shows that techniques such as Grad-CAM, LIME, SHAP values, and Saliency Maps can provide insights into DL credit decisions. The paper empirically validates its approach through performance comparisons, showing that 2D CNN models outperform 1D CNN models while maintaining interpretability. Then, the SHAP values were identified as the most effective explanation method, improving classification accuracy to 99% and 100% in test cases, and show the practical applicability of explainable DL models in financial decision-making. Then, the analysis that should be included in the review of DL applications in credit scoring [19] assesses whether DL architectures, specifically multilayer perceptron (MLP) networks, and DBNs, outperform conventional credit scoring models such as logistic regression, decision trees, random forests, and extreme gradient boosting (XGBoost). Using 10 retail credit scoring datasets, the paper analyses model performance based on 4 metrics: area under the curve (AUC), brier score (BS), partial Gini (PG), and expected maximum profit (EMP). The results show that XGBoost consistently outperforms all other models, achieving the highest ranking across all performance measures. However deep neural networks do not significantly outperform shallower ML models and are considerably more computationally expensive. The research further explores Bayesian statistical testing methods to compare classifiers, showing the advantages of Bayesian over frequentist hypothesis testing in credit scoring research. The paper concludes that DL is not the most suitable approach for credit scoring, and XGBoost should be preferred for optimal classification performance due to its superior accuracy and efficiency

The next section that should be analyzed is that ML and DL models are used to detect fraudulent activities by analyzing patterns and anomalies in transaction data. Techniques such as logistic regression, random forests, and decision trees have been used for evaluation in identifying fraudulent credit card transactions [1-3,8-9]. Among these, the random forest model showed good performance in predicting fraudulent activities [20]. The paper [21] highlights that random forest when combined with an ensemble feature selection technique integrating recursive feature elimination, information gain, and chi-squared ($\chi^2$) methods, improves fraud detection accuracy. The receiver operating characteristic score of 95.83% indicates the model's strong capability to distinguish fraudulent transactions from legitimate ones. Furthermore, the random forest model achieves an accuracy of 99.6%, with an F1-score of 99.6% and precision of 100%, showing that fraudulent transactions are correctly identified with minimal false positives. The research shows that the ability of the random forest to construct several decision trees and aggregate their results improves the model's generalization and resistance to overfitting. The ensemble feature selection approach also reduces training time while maintaining high model performance, making it highly efficient for real-time fraud detection. The paper [22] highlights that credit card fraud detection is a challenge due to the highly imbalanced nature of fraud datasets, where fraudulent transactions constitute a very small percentage of the total data. To address this issue, the study employs the synthetic minority over-sampling technique to balance the dataset. In this way, it improves the ability of the random forest model to detect fraudulent transactions accurately. The proposed model uses entropy as the guiding criterion for decision tree construction. It also provides optimal splits and reduces impurity in data classification. The model is trained on 1.3 million rows of training data and 550,000 rows of test data, representing both legitimate and fraudulent transactions. The research evaluates model performance using accuracy, precision, recall, F1-score, false negative rate, and false positive rate to assess its effectiveness. The results show that the random forest model achieves a high accuracy of 99.9% and reduces the number of undetected fraudulent transactions.

DL can process large data and provide good results in fraud detection. A systematic literature review highlighted the effectiveness of different DL models, such as CNN and LSTM networks, in detecting fraud across domains like credit card transactions and insurance claims [10]. These models have improved accuracy [3,23]. The CNN model is tested in the paper [24] on three benchmark credit card fraud datasets (European, Australian, and German datasets). The evaluation metrics used include the Matthews correlation coefficient, AUC, and cost of failure. The CNN model achieves an AUC of 87.64% on the European dataset and 70.56% on the German dataset, showing that can distinguish fraudulent transactions from legitimate ones. Furthermore, the CNN-based fraud detection model reduces the cost of failure, achieving a reduction of nearly 30% compared to traditional models. The finding of the paper [24] is that CNN outperforms traditional fraud detection methods, including random forest and SVM, by a margin of 5-10% in AUC scores. The CNN model is also found to be more resilient to imbalanced datasets, a common challenge in fraud detection, where fraudulent transactions account for less than 1% of total transactions. Then, the paper [25] addresses challenges in electricity theft detection including data imbalance, the curse of dimensionality, and the need for accurate classification. The proposed hybrid model integrates CNN for feature extraction and AdaBoost for classification. The model was trained and evaluated on the State Grid Corporation of China dataset, which contains 42,372 observations, including 3,579 fraudulent user records. The dataset was pre-processed using techniques like synthetic minority over-sampling technique to mitigate the data imbalance issue. Experimental results show that the CNN-AdaBoost model achieves a classification accuracy of 96.35%, outperforming traditional machine learning models such as logistic regression (84.05%), decision trees (90.05%), random forest (90.98%), and support vector machines (90.88%). The model also reduces the root mean square error to 0.2880 and achieves an F1-score of 95.60. Compared to standalone CNN and AdaBoost models, the hybrid approach improves fraud detection accuracy. The significant finding is in the paper [26] of graph convolutional networks (GCN) in detecting fraudulent activities and insider threats. The paper shows that traditional fraud detection models primarily rely on individual user properties and overlook structural relationships between users, which limits their accuracy in identifying coordinated fraudulent behaviors. GCN models overcome this limitation by transforming user behaviors and relationships into a graph-based structure, and the model learns both node-level (individual user) and edge-level (inter-user relationship) features. Unlike CNNs, which are effective for structured Euclidean data such as images, GCNs are designed to handle non-Euclidean data, ideal for fraud detection in complex networks. The paper analyses the GCN model using a real insider threat dataset (CMU CERT v4.2), which consists of 1,000 users and their digital activities over 16 months. The dataset includes event logs, such as email communication, logins, file transfers, and web activity, which are converted into graph representations. The research introduces a weighted adjacency matrix function that quantifies not only direct user connections but also behavioral similarities, improving the detection of fraudulent patterns. Experimental results show that GCN achieves a fraud detection accuracy of 94.5%, outperforming CNN (93%), random forest (85%), logistic regression (82%), and support vector machines (84%). Additionally, the recall rate of GCN is 83.3%, compared to an average recall of 70% for traditional models, meaning GCN reduces the number of undetected fraudulent activities.

Financial institutions employ clustering techniques to segment customers based on their behaviors and preferences, this is done in order to get personalized marketing strategies. AI personalization has been shown to build trust in digital finance with services that align with individual customer needs [27,28]. The paper [28] shows how collaborative filtering, content-based filtering, and hybrid recommendation models improve the accuracy of AI financial suggestions. The study evaluates AI in increasing user engagement, financial decision-making, and marketing strategies for financial institutions. Experimental results show that AI personalized payment recommendations increase user engagement by 27% compared to traditional, non-personalized recommendations. Additionally, fintech companies implementing AI personalization experience a 15% rise in customer retention rates and a 20% boost in conversion rates for financial products and services. An innovative approach was shown in the paper [29] where the evaluation of MyFinanceAI was presented. MyFinanceAI is an AI personal finance platform. The paper highlights how AI financial tools improve savings, but also financial literacy, and user engagement. It also addresses the growing complexity of modern financial management. MyFinanceAI system integrates DNNs and reinforcement learning to provide personalized financial recommendations based on real transaction data, spending habits, and financial goals.

A 6-month pilot study involving 1,000 users showed that 85% of users reported reduced financial stress, with a 43% decrease in financial anxiety scale scores. Additionally, monthly savings increased by 22%, translating to an average additional savings of $317 per user per month. The system's automated bill management feature also resulted in 92% of users avoiding late fees, leading to an average saving of $185 per user over the study period. The predictive budgeting module uses autoregressive integrated moving average and LSTM models, improving expense forecasting accuracy by 30% compared to traditional regression models. The findings also showed the impact of AI financial literacy tools, which resulted in a 40% increase in users' financial literacy scores over the study period. And the paper [30] highlights that chatbots and virtual assistants are offering customized financial recommendations, and automated support. The findings are that AI financial advisory systems analyze customer spending habits, investment preferences, and financial goals to deliver customized financial advice. Financial institutions that implement this personalization report a 20% increase in customer engagement and a 15% rise in customer retention rates, showing improved user experiences. The paper also highlights that AI chatbots and virtual assistants in finance reduce customer wait times by up to 35% and improve service accessibility. Additionally, AI models use natural language processing to improve financial security and fraud detection. AI chatbots assist in transactions, and fund transfers, leading to an 18% improvement in customer response times and a 25% reduction in operational costs for banks. These findings were analyzed and presented in Table 1.

*Table 1 - Comparative analysis of ML and DL models in financial applications*

| Model | Findings | Metrics and performance | Ref. |
|---|---|---|---|
| Decision Trees & SVM | Early ML models used for credit scoring and fraud detection; improved accuracy over statistical methods. | Improved early-stage ML classification in credit risk and fraud detection. | [8-9] |
| Neural networks | Improved default prediction in credit scoring; outperformed logistic regression in risk assessment. | Higher predictive power in financial applications compared to traditional ML models. | [13] |
| Deep Belief Networks (DBN) | Achieved higher accuracy than shallow networks in credit scoring; effective for deep feature extraction. | More accurate in capturing non-linear financial relationships than standard ML models. | [17] |
| Convolutional neural networks (CNN) | AUC of 87.64% in fraud detection (European dataset); cost of failure reduced by 30%. | Superior fraud detection performance in high-dimensional financial datasets. | [24] |
| Long short-term memory (LSTM) | Improved expense forecasting accuracy by 30%; improved financial planning using predictive models. | Better forecasting accuracy in financial risk assessment and management. | [29] |
| Random forest (RF) | 99.6% accuracy in fraud detection; 100% precision with ensemble feature selection, robust for real-time detection. | More effective in fraud detection with feature selection; minimized false negatives. | [20-22] |
| XGBoost | Outperformed all models in credit scoring; highest accuracy and efficiency across datasets. | Best performance in credit scoring; high interpretability and computational efficiency. | [19] |
| Graph convolutional networks (GCN) | Outperformed CNN in fraud detection (94.5% vs. 93%); improved recall rate by 10% using graph-based fraud analysis. | Better fraud detection than CNN by incorporating user relationship networks. | [26] |
| AI Personalization | Boosted customer retention by 15%, engagement by 27%, and conversion rates by 20% through AI-driven recommendations. | Higher engagement, retention, and improved financial advisory personalization. | [28-30] |
| Hybrid CNN-AdaBoost | 96.35% accuracy in electricity theft detection; improved fraud classification over standalone models. | Hybrid model achieved better fraud detection accuracy than standalone DL models. | [25] |

Regression analysis is often used in financial modeling, with quantitative tools in order to predict financial outcomes. The most common application is in stock price prediction. Accurately forecasting stock prices is needed in order to plan investment strategies. Regression models, both traditional and advanced, are used to analyze historical price and market indicators [31,32]. This approach presents a linear relationship between a stock's past performance and its future prices. With historical data, linear regression models can provide baseline predictions, though they may oversimplify complex market dynamics. The paper [31] analyses the linear regression in stock price prediction. The model is used to predict stock prices based on historical data, market indicators, macroeconomic factors, and sentiment analysis. In stock market forecasting, linear regression works by identifying trends and making predictions based on a line of best fit, minimizing the difference between actual and predicted values. The paper highlights that linear regression models have been effectively applied in predicting the closing stock prices of major companies. The research by Pahwa and Agarwal [33] showed the application of linear regression in forecasting the Google stock price using 14 years of historical data. The analyses found that linear regression provided high confidence values in trend forecasting, which is suitable for short-term stock price prediction. Another paper [34] further improved linear regression accuracy by incorporating principal component analysis (PCA) to select the most relevant financial indicators. The results showed that PCA linear regression models were more accurate than standalone models and even outperformed SVM in specific structured datasets. Techniques, such as LSTM networks, a type of recurrent neural network, have shown good capabilities in capturing temporal dependencies in stock price data. The paper [32] analyses LSTM to the Vietnamese stock market which achieved a 93% accuracy in predicting stock price trends, underscoring the model's effectiveness in handling sequential financial data. Combining traditional financial theories with machine learning models improves prediction accuracy. An example is integrating the Black-Scholes option pricing model with machine learning techniques to forecast stock prices in frontier markets [35]. The paper [36]presents option pricing with ANNs and the traditional Black-Scholes-Merton (BSM) model. It presents the introduction of a hybrid option pricing model (BS-ANN) and the comparative analysis of option pricing models, including Monte Carlo simulation, Heston, GARCH, Jump diffusion, and Cox-Ingersoll-Ross models. One of the findings is that the BS-ANN model outperforms traditional methods in terms of accuracy and stability. The paper analyzed the daily stock price data of Khodro, an automobile company, over the past year. The results showed that the BS-ANN model achieved the lowest standard deviation among all tested models, indicating higher precision in predicting European call and put option prices. Compared to standard Black-Scholes pricing, the hybrid ANN model reduced pricing errors and presented a more reliable alternative for real option pricing applications. The paper [37] examined the KOSPI200 options market using historical data from January to December 2023. The findings revealed that traditional models like the Black-Scholes and binomial tree model outperformed other methods in terms of accuracy. RMSE analysis demonstrated that Black-Scholes and the binomial tree method had the lowest prediction errors across all timeframes (weekly, monthly, and 50-day maturity). The Monte Carlo and variance reduction methods showed comparable performance, but their computational demands made them less practical for real-time trading. In contrast, ML models, particularly random forest and XGBoost, were less accurate in predicting option prices but excelled in identifying undervalued options. The investment simulation results highlighted that XGBoost and random forest achieved the highest profitability, showing that ML models could be useful for trading strategies focused on mispriced options. Significant results showed that Black-Scholes and the binomial tree had the lowest errors across all timeframes, 0.650 for weekly predictions, 0.641 for monthly, and 0.385 for 50-day maturities. Monte Carlo simulations with variance reduction techniques (AS, CV, IS) had slightly higher RMSE values, ranging between 0.485 and 0.686. ML models (XGBoost, Random Forest, MLP) had significantly higher RMSE values, with XGBoost at 6.307, random forest at 5.097, and MLP at 21.351 in weekly predictions.

It is important to emphasize that effective risk management is needed for financial stability. Regression analysis helps in quantifying and mitigating potential risks, specifically in volatility forecasting [38,39]. Predicting market volatility is important for risk assessment. ML models, including mnemonic DL techniques, have been developed to forecast midterm stock volatility, thereby helping in risk management and investment decisions [39]. Identifying anomalies in financial data is needed for potential risks. ML models can detect deviations from expected patterns, enabling proactive risk mitigation strategies. The paper [40] presents findings on financial anomaly detection, with DL models, gated recurrent units (GRU) with an attention mechanism (CA Module). The paper highlights traditional statistical models, such as distance-based outlier detection and density-based clustering approaches, with high-dimensional financial datasets. While methods like DBSCAN and k-NN have been used for anomaly detection, they suffer from high computational costs and parameter sensitivity when applied to large-scale market data. To overcome these challenges, the research introduces a hybrid GRU-CA model that improves anomaly detection. The GRU model, a variant of RNNs, addresses the vanishing gradient problem associated with traditional long-sequence financial data analysis. The results showed that GRU effectively retains long-term dependencies in financial time series data, making it superior to RNNs for market anomaly detection. The CA module improves the model's ability to filter relevant information by assigning different weights to market features, reducing noise, and increasing accuracy in detecting fraudulent or irregular trading patterns. The empirical evaluation was conducted using S&P 500 index data, where 70% of the dataset was used for training and 30% for testing. The RMSE metric was used to assess model performance. The findings show that the GRU model alone achieved an RMSE of 13.28, 13.27, and 13.29 across different prediction horizons, whereas the GRU-CA hybrid model significantly improved accuracy, reducing RMSE to 9.76, 9.78, and 9.74, demonstrating its superior ability to identify anomalies and estimate financial risks. On the other hand, the paper [41] presents anomaly detection in dynamic and evolving data environments. The research focuses on unsupervised anomaly detection methods and categorizes existing anomaly detection approaches into 3 main model-based categories: statistical-based models, clustering-based models, and nearest-neighbor models. The statistical-based approaches aim to model the normal behavior of a dataset and identify outliers as data points that significantly deviate from this normal distribution. The study highlights the Gaussian Mixture Model (GMM) as a commonly used statistical method, where an anomaly score is assigned based on the probability of data deviation from the estimated distribution. However, an underlying data distribution is required for effective anomaly detection, which limits its adaptability to evolving data streams.

The clustering-based approaches divide data into groups and classify anomalies based on their deviation from established clusters. BIRCH (balanced iterative reducing and clustering using hierarchies) is one of the earliest clustering-based anomaly detection techniques, designed to process large databases efficiently. The study also presents CluStream, which extends BIRCH by incorporating temporal characteristics of clusters for evolving data. Another significant advancement discussed is DenStream, a density-based clustering approach that improves on DBSCAN for detecting arbitrarily shaped clusters while handling noisy data streams. The DStream model further improves these methods by integrating grid-based clustering, significantly reducing computational overhead in anomaly detection. The nearest-neighbour-based techniques evaluate anomalies by measuring the distance between data points and their closest neighbours. k-Nearest Neighbors (k-NN) anomaly detection and Local Outlier Factor (LOF) are identified as highly effective for identifying localized anomalies in high-dimensional data. The study discusses incremental adaptations of these models for streaming data, such as iLOF (Incremental Local Outlier Factor), which updates local outlier scores dynamically as new data points arrive. However, memory and time complexity constraints pose challenges for real-time implementation. The comparative results found was presented in Table 2.

*Table 2 – Comparative analysis of ML and DL in option pricing and anomaly detection*

| Model | Findings | Metrics | Ref. |
|---|---|---|---|
| Linear Regression | Used for predicting stock prices based on historical data; high confidence values in trend forecasting; suitable for short-term predictions. | Applied to Google stock price prediction with 14 years of data; improved forecasting confidence but struggles with nonlinear trends. | [31, 33, 34] |
| LSTM (Long Short-Term Memory) | Achieved 93% accuracy in predicting Vietnamese stock prices; capable of capturing temporal dependencies for better long-term forecasting. | Outperformed regression models in stock price forecasting; superior handling of sequential financial data. | [32] |
| Black-Scholes-ANN Hybrid (BS-ANN) | Outperformed traditional Black-Scholes model in option pricing; lowest standard deviation and improved stability in European option pricing. | Achieved the lowest pricing error among all tested models; reduced mispricing risks compared to Monte Carlo and Black-Scholes. | [36] |
| Monte Carlo Simulation | Comparable performance to Black-Scholes but computationally expensive; variance reduction techniques improve accuracy. | RMSE ranged between 0.485-0.686; high accuracy but impractical for real-time trading. | [37] |
| XGBoost & Random Forest (RF) | Less accurate in option pricing but superior in detecting mispriced options; highest profitability in trading simulations. | XGBoost RMSE: 6.307; RF RMSE: 5.097; higher profitability but lower predictive accuracy. | [37] |
| GRU-CA Hybrid (Gated Recurrent Units + Contextual Attention) | Reduced RMSE from 13.28 to 9.74 in S&P 500 anomaly detection; captures long-term dependencies with attention-weighted feature selection. | RMSE reduced from 13.28 to 9.74; significantly better anomaly detection compared to standard RNNs. | [40] |
| Gaussian Mixture Model (GMM) | Assigns anomaly scores based on deviation from normal distribution; limited adaptability to evolving data streams. | Effective for financial anomaly detection; requires strong prior knowledge of data distribution. | [41] |
| Clustering-Based Models (BIRCH, DenStream, CluStream) | Improves computational efficiency for large datasets; captures evolving patterns in financial anomalies. | Reduces computational overhead; improves detection of financial anomalies in evolving datasets. | [41] |
| Nearest-Neighbor Models (k-NN, LOF, iLOF) | Effective in detecting localized anomalies; incremental updates allow adaptation to streaming data but computationally intensive. | Memory-intensive but provides high precision in anomaly detection; used in fraud detection and market monitoring. | [41] |

**Conclusion.** This systematic review analyzed the impact of ML and DL models in financial modeling, with a focus on classification and regression problems. The study explored ML/DL applications in credit scoring, fraud detection, algorithmic trading, stock price prediction, volatility forecasting, and option pricing, highlighting their strengths, limitations, and real applicability. A total of 41 papers were reviewed to assess emerging trends, model performance, and existing research gaps. Findings indicate that ML and DL models have improved financial decision-making by improving predictive accuracy, risk assessment, and trading efficiency. In credit scoring, DL models such as DBN and CNNs showed good performance over traditional ML models like logistic regression and decision trees. However, challenges such as fairness in credit decision-making and regulatory compliance remain concerns for widespread adoption. Fraud detection techniques using random forests, XGBoost, CNN, and hybrid models like CNN-AdaBoost achieved high detection accuracy. In stock price prediction, traditional linear regression models performed well for short-term forecasting but struggled with market complexities. LSTM networks outperformed traditional methods. In option pricing, hybrid models such as BS-ANN outperformed traditional pricing methods by reducing mispricing errors and improving stability. Monte Carlo simulations provided comparable results but were computationally expensive, limiting their real-time application. Anomaly detection remains an important area in financial risk management. Future research directions should focus on explainable AI methods to improve model transparency, federated learning to improve data privacy in financial applications, and quantum computing to address computational limitations in high-dimensional modelling tasks.

## References

1. Elizabeth Kuukua Woode Amoako, Victor Boateng, Ola Ajay, Tobias Kwame Adukpo, & Nicholas Mensah. (2025). Exploring the role of Machine Learning and Deep Learning in Anti-Money Laundering (AML) strategies within U.S. Financial Industry: A systematic review of implementation, effectiveness, and challenges. *Finance & Accounting Research Journal*, *7*(1), 22–36. https://doi.org/10.51594/farj.v7i1.1808.

2. Rajendran, S., John, A. A., Suhas, B., & Sahana, B. (2023). Role of ML and DL in detecting fraudulent transactions. In *Artificial intelligence for societal issues* (pp. 59–82). Springer International Publishing. https://doi.org/10.1007/978-3-031-12419-8_4.

3. Mienye, E., Jere, N., Obaido, G., Mienye, I. D., & Aruleba, K. (2024). Deep learning in finance: A survey of applications and techniques. *Ai*, *5*(4), 2066–2091. https://doi.org/10.3390/ai5040101.

4. Nazareth, N., & Ramana Reddy, Y. Y. (2023). Financial applications of machine learning: A literature review. *Expert Systems With Applications*, 119640. https://doi.org/10.1016/j.eswa.2023.119640.

5. Zheng, Y., Xu, Z., & Xiao, A. (2023). Deep learning in economics: A systematic and critical review. *Artificial Intelligence Review*. https://doi.org/10.1007/s10462-022-10272-8.

6. Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. *Applied Soft Computing*, *90*, 106181. https://doi.org/10.1016/j.asoc.2020.106181.

7. Shi, S., Tse, R., Luo, W., D'Addona, S., & Pau, G. (2022). Machine learning-driven credit risk: A systemic review. *Neural Computing and Applications*. https://doi.org/10.1007/s00521-022-07472-2.

8. Sahin, Y., & Duman, E. (2011, March). Detecting credit card fraud by decision trees and support vector machines. In *Proceedings of the international multiconference of engineers and computer scientists* (Vol. 1, pp. 1-6). https://www.iaeng.org/publication/IMECS2011/IMECS2011_pp442-447.pdf.

9. Huang, C.-L., Chen, M.-C., & Wang, C.-J. (2007). Credit scoring with a data mining approach based on support vector machines. *Expert Systems With Applications*, *33*(4), 847–856. https://doi.org/10.1016/j.eswa.2006.07.007.

10. Peivandizadeh, A., Hatami, S., Nakhjavani, A., Khoshsima, L., Qazani, M. R. C., Haleem, M., & Alizadehsani, R. (2024). Stock market prediction with transductive long short-term memory and social media sentiment analysis. *IEEE Access*, 1. https://doi.org/10.1109/access.2024.3399548.

11. Khalil, F., & Pipa, G. (2021). Is deep-learning and natural language processing transcending the financial forecasting? Investigation through lens of news analytic process. *Computational Economics*. https://doi.org/10.1007/s10614-021-10145-2.

12. Liu, C., Arulappan, A., Naha, R., Mahanti, A., Kamruzzaman, J., & Ra, I.-H. (2024). Large language models and sentiment analysis in financial markets: A review, datasets and case study. *IEEE Access*, 1. https://doi.org/10.1109/access.2024.3445413.

13. Hang, V., Sivakulasingam, S., Wang, H., Wong, S. T., Ganatra, M. A., & Luo, J. (2024). Credit risk prediction using machine learning and deep learning: A study on credit card customers. *Risks*, *12*(11), 174. https://doi.org/10.3390/risks12110174.

14. Kumar, A., Sharma, S., & Mahdavi, M. (2021). Machine learning (ML) technologies for digital credit scoring in rural finance: A literature review. *Risks*, *9*(11), 192. https://doi.org/10.3390/risks9110192.

15. Kozodoi, N., Jacob, J., & Lessmann, S. (2021). Fairness in credit scoring: Assessment, implementation and profit implications. *European Journal of Operational Research*. *297*(3), 1083-1094. https://doi.org/10.1016/j.ejor.2021.06.023.

16. Albanesi, S., & Vamossy, D. F. (2019). *Predicting consumer default: A deep learning approach* (No. w26165). National Bureau of Economic Research.

17. Hayashi, Y. (2022). Emerging trends in deep learning for credit scoring: A review. *Electronics*, *11*(19), 3181. https://doi.org/10.3390/electronics11193181.

18. Dastile, X., & Celik, T. (2021). Making deep learning-based predictions for credit scoring explainable. *IEEE Access*, 1. https://doi.org/10.1109/access.2021.3068854.

19. Gunnarsson, B. R., Vanden Broucke, S., Baesens, B., Óskarsdóttir, M., & Lemahieu, W. (2021). Deep learning for credit scoring: Do or don't? *European Journal of Operational Research*, *295*(1), 292–305. https://doi.org/10.1016/j.ejor.2021.03.006.

20. Afriyie, J. K., Tawiah, K., Pels, W. A., Addai-Henne, S., Dwamena, H. A., Owiredu, E. O., Ayeh, S. A., & Eshun, J. (2023). A supervised machine learning algorithm for detecting and predicting fraud in credit card transactions. *Decision Analytics Journal*, *6*, 100163. https://doi.org/10.1016/j.dajour.2023.100163.

21. Akazue, M. I., Debekeme, I. A., Edje, A. E., Asuai, C., & Osame, U. J. (2023). UNMASKING FRAUDSTERS: Ensemble features selection to enhance random forest fraud detection. *Journal of Computing Theories and Applications*, *1*(2), 201–211. https://doi.org/10.33633/jcta.v1i2.9462.

22. Mihali, S.-I., & Niță, Ș.-L. (2024). Credit card fraud detection based on random forest model. In *2024 international conference on development and application systems (DAS)*. IEEE. https://doi.org/10.1109/das61944.2024.10541240.

23. Chen, Y., Zhao, C., Xu, Y., & Nie, C. (2025). Year-over-Year Developments in Financial Fraud Detection via Deep Learning: A Systematic Literature Review. http://dx.doi.org/10.48550/arXiv.2502.00201.

24. Raghavan, P., & Gayar, N. E. (2019). Fraud Detection using Machine Learning and Deep Learning. In *2019 International conference on computational intelligence and knowledge economy (ICCIKE). IEEE*. https://doi.org/10.1109/iccike47802.2019.9004231.

25. Nirmal, S., Patil, P., & Kumar, J. R. R. (2024). CNN-AdaBoost based hybrid model for electricity theft detection in smart grid. *E-Prime - Advances in Electrical Engineering, Electronics and Energy*, *7*, 100452. https://doi.org/10.1016/j.prime.2024.100452.

26. Jiang, J., Chen, J., Gu, T., Choo, K.-K. R., Liu, C., Yu, M., Huang, W., & Mohapatra, P. (2019). Anomaly detection with graph convolutional networks for insider threat and fraud detection. In *MILCOM 2019 - 2019 IEEE military communications conference (MILCOM)*. IEEE. https://doi.org/10.1109/milcom47813.2019.9020760

27. Kanaparthi, V. (2024). AI-based personalization and trust in digital finance. https://doi.org/10.48550/arXiv.2401.15700.

28. Abba, S. (2022). Ai in Fintech: Personalized Payment Recommendations for Enhanced User Engagement. *Technology (IJRCAIT)*, *5*(01), 13-20. https://surl.li/tfxaxv.

29. Talasila, S. D. (2024). AI-Driven Personal Finance Management: Revolutionizing Budgeting and Financial Planning. *International Research Journal of Engineering and Technology*, *11*(7), 397-400. https://surl.li/dqquaj.

30. Bhuiyan, M. S. (2024). The role of ai-enhanced personalization in customer experiences. *Journal of Computer Science and Technology Studies*, *6*(1), 162–169. https://doi.org/10.32996/jcsts.2024.6.1.17.

31. Soni, P., Tewari, Y., & Krishnan, D. (2022). Machine learning approaches in stock price prediction: A systematic review. *Journal of Physics: Conference Series*, *2161*(1), 012065. https://doi.org/10.1088/1742-6596/2161/1/012065.

32. Phuoc, T., Anh, P. T. K., Tam, P. H., & Nguyen, C. V. (2024). Applying machine learning algorithms to predict the stock price trend in the stock market – The case of Vietnam. *Humanities and Social Sciences Communications*, *11*(1). https://doi.org/10.1057/s41599-024-02807-x.

33. Pahwa, K., & Agarwal, N. (2019). Stock market analysis using supervised machine learning. In *2019 international conference on machine learning, big data, cloud and parallel computing (comitcon)*. IEEE. https://doi.org/10.1109/comitcon.2019.8862225.

34. Misra, M., Yadav, A. P., & Kaur, H. (2018). Stock market prediction using machine learning algorithms: A classification study. In *2018 international conference on recent innovations in electrical, electronics & communication engineering (ICRIEECE)*. IEEE. https://doi.org/10.1109/icrieece44171.2018.9009178.

35. Chowdhury, R., Mahdy, M. R. C., Alam, T. N., Al Quaderi, G. D., & Arifur Rahman, M. (2020). Predicting the stock price of frontier markets using machine learning and modified Black–Scholes Option pricing model. *Physica A: Statistical Mechanics and Its Applications*, *555*, 124444. https://doi.org/10.1016/j.physa.2020.124444

36. Shahvaroughi Farahani, M., Babaei, S., & Esfahani, A. (2024). "Black-Scholes-Artificial Neural Network": A novel option pricing model. *International Journal of Financial, Accounting, and Management*, *5*(4), 475–509. https://doi.org/10.35912/ijfam.v5i4.1684.

37. Kim, S., Kim, J., & Song, J. (2024). Option pricing and profitability: A comprehensive examination of machine learning, Black-Scholes, and Monte Carlo method. *Communications for Statistical Applications and Methods*, *31*(5), 585–599. https://doi.org/10.29220/csam.2024.31.5.585.

38. Stanković, Z. Z., Rajic, M. N., Božić, Z., Milosavljević, P., Păcurar, A., Borzan, C., Păcurar, R., & Sabău, E. (2024). The volatility dynamics of prices in the european power markets during the COVID-19 pandemic period. *Sustainability*, *16*(6), 2426. https://doi.org/10.3390/su16062426.

39. Li, X., Li, Y., Liu, X. Y., & Wang, C. D. (2019). Risk management via anomaly circumvent: mnemonic deep learning for midterm stock prediction. https://doi.org/10.48550/arXiv.1908.01112.

40. Wang, B., Dong, Y., Yao, J., Qin, H., & Wang, J. (2024). Exploring anomaly detection and risk assessment in financial markets using deep neural networks. *International Journal of Innovative Research in Computer Science and Technology*, *12*(4), 92–98. https://doi.org/10.55524/ijircst.2024.12.4.15.

41. Salehi, M., & Rashidi, L. (2018). A survey on anomaly detection in evolving data. *ACM SIGKDD Explorations Newsletter*, *20*(1), 13–23. https://doi.org/10.1145/3229329.3229332.

Отримано 20.03.2025

*УДК 004.8:336.76:519.2*

***Микола Миколайович Злобін[1], Володимир Маркович Базилевич[2]***

[1]аспірант

Національний університет «Чернігівська політехніка» (Чернігів, Україна)

**E-mail:** mykolay.zlobin@gmail.com. **ORCID:** https://orcid.org/0009-0000-7653-6109. **Scopus Author ID:** 59337918100

[2]кандидат економічних наук, доцент кафедри Інформаційних та комп'ютерних наук, директор ННІ ЕІТ

Національний університет «Чернігівська політехніка» (Чернігів, Україна)

**E-mail:** bazvlamar@stu.cn.ua. **ORCID:** https://orcid.org/0000-0001-8935-446X. **Scopus Author ID:** 57214432127

# СИСТЕМАТИЧНИЙ ОГЛЯД ГЛИБОКОГО ТА МАШИННОГО НАВЧАННЯ ДЛЯ ФІНАНСОВОГО МОДЕЛЮВАННЯ

*Машинне навчання (ML) і глибоке навчання (DL) покращують фінансову аналітику, зокрема кредитний скоринг, виявлення шахрайства, прогнозування цін на акції та ціноутворення опціонів. Традиційні методи не справляються з великими масивами даних і складністю ринків, що робить ML/DL необхідними для підвищення точності та оцінки ризиків. У цьому огляді проаналізовано найсучасніші програми ML/DL, порівняно методології, ефективність та проблеми, такі як інтерпретованість моделей, якість даних та концептуальний дрейф. У цій статті розглядаються застосування ML та DL для класифікації та регресійних задач у фінансовій сфері, оцінюються їхні методології, ефективність та проблеми. Хоча попередні дослідження підкреслили ефективність мереж глибокого переконання (DBN), згорткових нейронних мереж (CNN) і мереж з довгою короткочасною пам'яттю (LSTM), невирішеними залишаються питання інтерпретованості моделей, якості даних і адаптивності до фінансових ринків, що розвиваються. Багато моделей ML та DL функціонують як «чорні скриньки», що обмежує їхню прозорість для регульованих фінансових середовищ. Крім того, фінансові дані страждають від незбалансованого розподілу та шуму, що впливає на точність моделей. Проблема концептуального дрейфу також зберігається, оскільки ринкові умови змінюються з часом, що робить статичні моделі менш надійними. Результати показують, що моделі DL перевершують традиційні підходи.*

*CNN та DBN покращують кредитний скоринг, а випадкові ліси та XGBoost досягають 99,6% точності у виявленні шахрайства. LSTM-мережі досягають 93 % точності у прогнозуванні цін на акції, перевершуючи регресійні моделі. У ціноутворенні опціонів гібриди Black-Scholes-ANN покращують стабільність і точність порівняно з методами Монте-Карло. Моделі GRU-CA покращують виявлення аномалій, зменшуючи RMSE з 13,28 до 9,74, а GCN перевершують CNN, досягаючи 94,5% точності виявлення шахрайства. Залишаються проблеми з пояснюваністю, цілісністю даних і адаптивністю до ринку. Майбутні дослідження мають бути зосереджені на пояснюваному ШІ, федеративному навчанні та гібридних моделях ML-економетрії для підвищення точності, прозорості та реальної застосовності у прийнятті фінансових рішень.*

***Ключові слова:*** *глибоке навчання; машинне навчання; проблеми класифікації; регресійні моделі; фінансова аналітика.*

*Таблиця: 2. Бібл.: 41.*

Zlobin, M., Bazylevych, V. (2025). Systematic review of deep and machine learning for financial modeling. *Technical sciences and technologies*, (1(39)), 184-195.