Grenze International Journal of Engineering and Technology, June Issue

Survey Paper on Understanding Financial Markets Apps using Machine Learning

Ashoka D V[1], Deepa Konnur[2], Harshitha V[3], Chaitra S P[4], Abhay S K[5] and Anirudh A[6]
[1]Professor, Department of Information Science and Engineering, JSS Academy of Technical Education, Bengaluru, India
Email: dr.dvashoka@gmail.com
[2]Asst. Professor, Department of Information Science and Engineering, JSS Academy of Technical Education, Bengaluru, India
Email: deepakonnur@jssateb.ac.in
[3-6]Department of Information Science and Engineering, JSS Academy of Technical Education, Bengaluru, India
Email: harshithav09334@gmail.com, chaitrasp073@gmail.com, abhayskulkarni11@gmail.com, anirudhashrit97@gmail.com

***Abstract*— The financial markets are multifaceted and rapidly evolving, presenting challenges such as fragmented data access, inefficient tools for portfolio management, and inadequate predictive insights. With advancements in technology, particularly machine leaming, there is immense potential to address these issues by consolidating features and providing actionable insights. Our project introduces the All-in-One Financial Markets App, which leverages cutting-edge machine learning techniques to integrate essential functionalities such as stock price visualization, sentiment analysis, portfolio tracking, and trading simulations. Stock price visualization enables users to view historical trends and identify patterns using intuitive graphical representations. Sentiment analysis processes financial news und social media data to gauge market sentiment, giving users a predictive edge. Portfolio tracking consolidates investments, offering detailed metrics like ROI and diversification analysis. The app's trading simulation allows both novice and experienced investors to back test strategies with historical data, fostering a deeper understanding of market dynamics.**

Additionally, real-time data integration ensures users are continuously informed about market movements, while the machine learning models deliver precise forecasts and actionable recommendations. By combining these features, the application aims to eliminate inefficiencies, reduce complexity, and provide a seamless, user-friendly experience for all types of investors. In essence, this app serves as a comprehensive tool to empower users with informed decision-making capabilities, making financial markets accessible and manageable.

***Keywords*— Financial Markets, Machine Learning. Portfolio Management, Sentiment Analysis, Real-Time Analytics. Index Terms—component, formatting, style, styling, insert.**

I. INTRODUCTION

The financial market is a crucial element of global economies, significantly impacting the monetary well-being of consumers, traders, and financial institutions. In particular, stock trading involves a substantial portion of the population, with surveys indicating that one in seven people in China participate in it, highlighting the growing role of stocks in the global economy.

Grenze ID: 01.GIJET.12.2.134_21

[1] Machine learning algorithms are attractive for financial time series forecasting due to their ability to approximate non-linear functions, handle noisy, non-stationary data, and discover latent patterns [4]. Despite ML's growing importance in finance, academic consolidation and standardisation in this field remain sparse. Existing literature lacks a compelling analysis of different algorithms and their findings. This study seeks to address this gap by performing a detailed and systematic review of prior research on trading algorithms and by proposing strong, evidence-based hypotheses regarding algorithm performance using rank-based analytical methods. By gathering many samples from different experimental methodologies, the study avoids biases from varying financial interfaces and datasets, aiming to represent true differences between algorithm classes. Recurrent networks are specifically highlighted for their potential to model temporal dynamics and long-term dependencies within time series. The stock market is influenced by numerous macroeconomic factors, including political events, firm policies, general economic conditions, investor expectations, institutional investor decisions, other stock market movements, and investor psychology. Many empirical studies have focused on predicting stock index movement in developed financial markets [6]. Artificial Intelligence (AI) and Machine Learning (ML) have increasingly shaped the financial industry as advancements in computational power and algorithmic design have progressed. Although the earliest uses of AI and ML in finance emerged in the 1980s, their application has since expanded from basic analytical functions to sophisticated activities such as stock price prediction, anomaly detection, and automated trading. These technologies are transforming financial operations by supporting data-driven decision-making, streamlining routine processes, and enabling personalized financial services. According to the World Economic Forum (2018), the integration of AI into financial systems could contribute economic value exceeding one trillion dollars by 2025.

II. Research Methodology

"Machine Learning in Financial Market Forecasting" highlights how ML is transforming the financial industry. Core applications include portfolio management, stock price prediction, algorithmic trading, risk management, fraud detection, credit scoring, customer service, and financial time series analysis. To achieve these, a variety of techniques are used, ranging from traditional ML models and ensemble methods to advanced deep learning, data augmentation, feature extraction, and sentiment analysis. However, challenges remain, such as limited availability of long-term financial data, the black-box nature of models, risks of overfitting, and high computational costs. Despite these limitations, the benefits are substantial—machine learning improves efficiency, enhances accuracy, uncovers hidden patterns in data, enables automated quantitative analysis, and adapts quickly to dynamic market conditions, making it a powerful tool in modern financial forecasting.

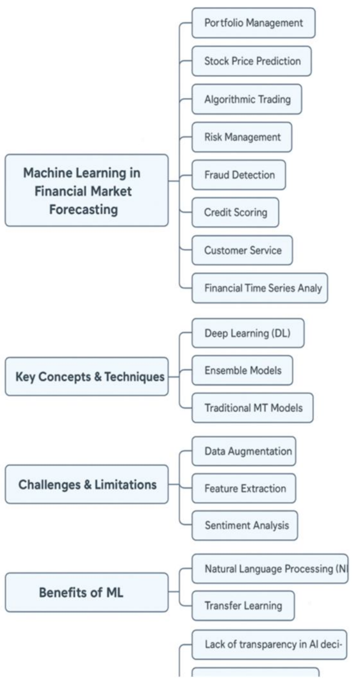

Fig. 1. System Design and Methodology for All-in-One Financial Markets App Using Machine Learning

III. ANALYSIS OF EXISTING METHODOLOGIES

The different approaches are discovered through different models.SK-GCN model delivers excellent results in stock classification, achieving significantly higher accuracy (16%-25%) higher than other models, particularly in small sample training. This model addresses the limitations of traditional manual classification and can be used to classify equities into new concept sectors with high accuracy even with a small number of labels[1].The application of advanced machine learning approaches, such as text data analytics and ensemble methods, has greatly increased prediction accuracies .Improved Model Performance with While social media data alone doesn't perform better than market data and technical indicators, combining data from textual sources with market data and technical indicators significantly enhances model performance and prediction accuracies[2].Inherent Complexity of Stock Market Data: SMP remains one of the most challenging research areas due to the dynamic, erratic, chaotic, non-linear, stochastic, and unreliable nature of stock market data.ML is an ideal candidate for empirical discovery due to its ability to identify complex patterns, break down high-dimensional data, and detect patterns with minimal supervision (unsupervised methods). It can be used to discover empirical relationships from data (e.g., clustering, autoencoders) that fit into financial-economic theories, aiding theory construction. Progress in model interpretability helps in this context [3]. machine learning (ML) algorithms generally outperform most traditional stochastic methods in financial market forecasting, with recurrent neural network (ANNR) demonstrating superior performance over feed-forward neural networks and support vector machines, suggesting the presence of exploitable temporal dependencies. However, the field is hampered by a significant lack of academic consolidation and standardization, including varied testing metrics, datasets, and optimization details, making direct parametric comparisons across studies unfeasible due to inherent heterogeneity. Further limitations include the "black-box" nature of ML models, which compromises decision transparency, and the problematic use of accuracy as a performance metric without sufficient context on profitability, alongside a potential for researcher bias and data manipulation in the absence of standardized input data [4]. Despite the advancements, significant limitations persist, as no universal solution exists for consistently accurate stock price prediction. The efficacy of these models is highly dependent on meticulous hyperparameter tuning, which is a crucial and complex stage in maximizing performance, and models can fail if not efficiently trained with fresh data. Furthermore, the dynamic and evolving nature of the stock market, influenced by various factors, means that current approaches may see a degradation in performance over time, as underlying challenges constantly change [5]. Nifty50 Index prediction revealed that Linear Regression (LR) and Artificial Neural Networks (ANN) delivered consistently similar and strong performance across varying dataset sizes, though ANN often required more training time. Stochastic Gradient Descent(SGD) proved more effective than Support Vector Machine (SVM) with larger datasets, suggesting its better scalability. However, a significant limitation was the poor performance and severe overfitting observed in Adaptive Boost (AdaBoost), Random Forest (RF), k-Nearest Neighbors (kNN), and Decision Trees (DT), especially as dataset size increased. These models often showed negative R2 values on test data and failed to predict accurately beyond an initial period, indicating their unsuitability for dynamic stock market forecasting with extensive historical data. Ultimately, the research underscores the inherent difficulty and complex nature of stock market prediction, making it challenging for any single model to provide consistently accurate forecasts [6].High implementation costs, limited data infrastructure, and complex regulatory requirements further constrain the deployment of advanced financial models. Additionally, ethical issues—including data privacy, algorithmic bias, fairness, transparency, and accountability—remain critical concerns. These challenges are compounded by systemic risks such as potential market disruptions caused by AI-driven trading and the possibility of workforce displacement due to automation [7]. Despite these advancements, the study highlights several limitations, including many unresolved concerns and difficulties that indicate room for development in ML approaches for currency prediction. A significant challenge is the heterogeneity of evaluation methodologies and metrics across studies, making direct comparisons unfeasible and necessitating the normalisation of effect sizes for meta-analysis. Furthermore, the common practice of using a percentage split for dataset validation is considered "suspect", with a recommendation for a more robust three-part division into training, test, and validation sets. The inherent limitations of meta-analyses also mean they cannot compensate for poor study design or bias in individual research [8]. The research highlights several limitations: the reliance solely on historical closing prices may restrict the model's comprehensiveness, and the exclusive use of the Bombay Stock Exchange (BSE) dataset limits the generalisability of the findings. Future work should integrate additional factors like news and investor sentiment and incorporate diverse stock market datasets to enhance robustness and accuracy. Additionally, while hybrid models address the limitations of individual linear (ARIMA) and non-linear (LS-SVM) models, they still face challenges like interpretability and computational cost in some complex scenarios [9]. The research acknowledges the inherent difficulty of increased operational burden and transaction costs in medium to long-term management, especially due to rapid fluctuations in abnormal returns causing frequent reversing trades. Other limitations include the “factor zoo” problem in conventional multifactor models, where distinguishing useful factors is challenging, and the potential for overlearning in fine-tuning due to limited data samples. In the FX market, the restricted number of G10 currency pairs makes it challenging to adequately spread risk when constructing portfolios [10]. The InvestingIQ platform successfully provides retail investors with real-time stock market information, financial performance metrics, and public sentiment analysis by leveraging machine learning and natural language processing. It incorporates Facebook Prophet and Random Forest models to offer independent future stock price predictions, which are stated to achieve up to 50%accuracy. However, a key limitation identified is the need to improve the model’s accuracy and the potential to feed more training parameters into the system for enhanced performance. The inherently noisy and stagnant nature of stock market data also presents a general challenge to accurate prediction [11]. Long Short-Term Memory (LSTM) models significantly outperformed Linear Regression (LR) and Convolutional Neural Network (CNN) models in predicting Apple stock prices across one, five, and ten-year timeframes, showing higher R-squared scores and lower error rates, primarily due to LSTM’s ability to capture long-term data relationships. LSTM also achieved high accuracy when tested with other stock variables like open, high, and low prices. However, the research highlights several limitations in financial time series prediction, including the inherent unpredictability of financial markets influenced by various economic and geopolitical factors. Furthermore, complex deep learning models, despite their capabilities, lack interpretability, hindering their practical adoption. Other challenges include their tendency to overfit and struggle with new data, their heavy reliance on historical trends, substantial data requirements, and sensitivity to market volatility, alongside the need for considerable computational power. LSTM’s performance, specifically, can also be affected by feature selection and hyperparameter tuning. [12]. These models suffer from limited interpretability and demand substantial computational resources for both training and fine-tuning. Furthermore, their performance is highly susceptible to hyperparameter selection, where incorrect values can lead to suboptimal predictions, and the inherent unpredictability of financial markets, influenced by external variables and geopolitical events, makes consistently accurate forecasting challenging even with advanced models [13]. Monte Carlo simulation was identified as the most accurate for stock price prediction, capable of providing a share price forecast by considering a variety of scenarios. The paper suggests that algorithm trading, which leverages machine learning, offers significant advantages such as lower operating costs, reduced reliance on human emotions, and improved accuracy, making it more accessible to the general public. risk of data loss due to hacker access as more personal and financial information becomes available online. A significant challenge is that the wealth of information available online about companies is often insufficient for precise share price anticipation, primarily due to the inherent uncertainty of the market. Human emotions like greed and fear, leading investors to make decisions based on guesses rather than fundamental analysis, can cause market movements that existing models may fail to understand, potentially leading to inaccurate predictions [14].

TABEL I: SUMMARY OF RESEARCH PAPERS ON FINANCIAL TIME SERIES PREDICTION

S. No	Year	Authors	Title	Key Findings	Methodologies	Advantages	Limitations
1	2023	Huize Xu et al	Development of Financial Stock Classification Using Graph Convolutional Neural Networks	Developed a Graph Convolutional Neural Network (GCN) model called SK-GCN for financial stock classification	Graph Convolutional Neural Networks (GCN): Used for extracting and learning from the relationships among stocks.	High classification performance using GCN compared to traditional models. Data-driven approach: Automates and improves accuracy in a complex and noisy stock environment.	Focused on Chinese stock markets — results may not generalize globally. Depends on quality and structure of financial data; noisy or incomplete data can affect performance.
2	2021	Oladayo Bello et al.	Design and Implementation of Electronic Voting System Using	The study demonstrates that integrating smart contracts on blockchain	Front-end interface using webtechnologies (HTML/CSS/JS) Back-end using	Transparency: Immutable records ensure trustworthy results.	Technical complexity requiring understanding of blockchain, smart

			Smart Contract on Blockchain	significantly increases voting security, transparency, and integrity	Remix IDE, Metamask, and Ganache for simulating the blockchain network	Security: Reduced risk of vote tampering due to blockchain's cryptographic features.	contracts, and cryptocurrency wallets.
3	2022	Kristof Lommers et al.	Confronting Machine Learning With Financial Research	Machine learning holds great potential in empirical finance, including tasks such as estimation, empirical discovery, hypothesis testing, causal inference, and prediction	estimation, empirical discovery, hypothesis testing, causal inference, and prediction.	Enables high-dimensional pattern recognition without strong prior assumptions. Offers flexibility and generalization capabilities for discovering relationships in large datasets.	Financial data often violates Assumptions common in ML (e.g., stationarity, i.i.d. samples). Misuse or blind application of ML in finance can lead to misleading or non-parsimonious models.
4	2019	Lukas Ryll et al.	Evaluating the Performance of Machine Learning Algorithms in Financial Market Forecasting: A Comprehensive Survey	Recurrent Neural Networks (RNNs), especially LSTM and GRU, outperform feed-forward neural networks and support vector machines.	A meta-analysis of over 150 studies covering 225 experiments and 2085 performance values. Focus on non-parametric rank analysis to determine algorithm performance.	Provides strong evidence of ML superiority over traditional methods in many scenarios. Demonstrates the value of RNNs and deep learning for capturing complex financial patterns.	Performance metrics like accuracy may not reflect real-world profitability. ML models remain black-boxes, limiting interpretability in critical financial decisions.
5	2023	Gaurang Sonkavde et al	Forecasting Stock Market Prices Using Machine Learning and Deep Learning Models	Ensemble models (particularly Random Forest + XG-Boost + LSTM) significantly improve accuracy and forecasting reliability for stock price prediction.	Machine Learning: Linear Regression, SVM, Naive Bayes, KNN, Logistic Regression.	Combines ML/DL insights from many studies into a unified performance framework. Demonstrates improved results via hybrid ensemble methods.	Model performance is highly sensitive to hyperparameters; default values often underperform. High computational cost for ensemble models, especially deep networks.
6	2024	Agampreet Saini et al	Financial Time Series Prediction on Apple Stocks using Machine and Deep Learning Models	LSTM demonstrated superior performance in predicting stock prices compared to traditional models like Linear Regression and CNN	Use of statistical metrics such as Mean AbsoluteError (M AE), Root Mean Squared Error (RMSE), and R-squared Score to evaluate model performance.	The approach allows for a comprehensive understanding of the capabilities and limitations of different predictive models.	The reliance on historical data may not fully capture future market dynamics, leading to potential inaccuracies in predictions.
7	2023	Michael Ayitey Junior, et al	Forex Market Forecasting Using Machine Learning:	LSTM and Artificial Neural Networks (ANNs) are the most	Algorithms studied: LSTM, GRU, SVM, ANN,	One of the most comprehensive reviews	Forex market remains highly volatile and nonlinear, making

			Systematic Literature Review and Meta-analysis	frequently used and effective models for predicting Forex market movements.	RNN, FLANN, HMM, SVR, and hybrid models. Evaluation Techniques: Validation methods like k-fold cross-validation, percentage split, etc.	focused exclusively on ML models for the Forex market. Identifies best-performing models and common pitfalls in Forex prediction.	prediction challenging. Many studies lack standardized evaluation metrics and reproducible experimental setups.
8	2022	Dr. Gurjeet Singh et	Machine Learning Models in Stock Market Prediction	Linear Regression (LR) and Artificial Neural Network (ANN) consistently performed well across all datasets. As data size increases, Stochastic Gradient Descent (SGD) outperformed Support Vector Machine (SVM).	Historical data of Nifty 50 Index (from 22 April 1996 to 16 April 2021) was divided into 4 subsets (12.5%, 25%, 50%, and 100%).	Longitudinal dataset (25 years) provides robustness against market volatility. Use of multiple performance metrics increases reliability of model evaluation. Cross-validation strengthens confidence in results.	Ensemble models like AdaBoost and RF overfitted on large datasets. ANN took longer time to train compared to simpler models like LR.
9	2024	Jyotirmayee Behera et al	An approach to portfolio optimization with time series forecasting algorithms and machine learning techniques	Improved decision-making for investors through better stock price predictions. Effective risk management capabilities by capturing both linear and nonlinear patterns.	Combination of ARIMA and LS-SVM for stock selection. Utilization of historical data from the Bombay Stock Exchange (BSE) for validation.	Enhanced forecasting accuracy by integrating linear (ARIMA) and nonlinear (LS-SVM) models. Demonstrated superior performance in portfolio optimization	Reliance on historical closing prices may limit the model's comprehensiveness. The study is based solely on the BSE dataset, which may restrict generalizability.
10	2023	Kazuki Amagai et al	Long-Term Modelling of Financial Machine Learning for Active Portfolio Management	Combining learning data from various time scales to improve model performance in long-term financial predictions. Sequentially relearning the autoencoder to adapt to dynamic	Quantile Portfolio Construction: Grouping stocks into quantiles based on abnormal returns for portfolio management.	Improved generalization in long-term modelling via data augmentation.	Decline in learning precision with longer time scales due to reduced data availability. Model complexity can hinder interpretability

IV. CONCLUSION

Machine learning (ML) models, while promising for stock market prediction, face significant limitations stemming from the inherent complexity of the market itself and the models' own characteristics. Despite impressive results from specific models like SK-GCN and recurrent neural networks (RNNs), no single model provides a universal or consistently accurate solution. The stock market's dynamic, erratic, and nonlinear nature is the primary obstacle. It's influenced by a multitude of unpredictable factors like political events and human emotions, which can lead to rapid fluctuations that are difficult for models to anticipate. Models are heavily reliant on data quality, and the available data can be noisy, stagnant, or insufficient, leading to poor model performance. The exclusive use of historical data, like closing prices, can limit a model's comprehensiveness and generalizability. Certain models, like AdaBoost, Random Forest, and KNN, suffer from severe overfitting and poor performance as the dataset size increases, making them unsuitable for forecasting with extensive historical data. Findings from one dataset, such as the Bombay Stock Exchange (BSE), may not be generalizable to other markets, highlighting a need for more diverse datasets in research. Even advanced models have their own set of challenges, preventing them from being a flawless solution. Many deep learning models, like LSTMs, are difficult to interpret. This lack of transparency makes it hard for investors to understand the reasoning behind a prediction, which can compromise trust and practical adoption. A common problem, especially with complex deep learning models, is overfitting to historical data. This leads to models that perform well on past data but fail to generalize to new, unseen market conditions. Training and finetuning complex models often require significant computational resources, which can be a barrier to implementation. Model efficacy is highly dependent on meticulous hyperparameter tuning, a crucial yet complex step. Suboptimal predictions can result from incorrect hyperparameter values. There is a significant lack of academic consolidation and standardization. Varied testing metrics, datasets, and optimization details make it difficult to compare the performance of different models directly. The use of a simple percentage split for dataset validation is considered “suspect” and less robust than a three part division (training, test, and validation sets). The application of AI in finance raises pressing ethical issues, including data privacy, algorithmic bias, fairness, and transparency. Systemic risks, such as market disruptions from AI-driven trading, and potential job displacement due to automation are also significant concerns. Simply measuring accuracy is an insufficient performance metric without sufficient context on profitability and risk, as it doesn’t account for transaction costs or risk-to reward ratios.

REFERENCES

[1] H. Xu, Y. Zhang and Y. Xu, "Promoting Financial Market Development-Financial Stock Classification Using Graph Convolutional Neural Networks," in IEEE Access, vol. 11, pp. 49289-49299, 2023, doi: 10.1109/ACCESS.2023.3275085.

[2] D. Yin, M. Zhang, B. Wei and W. Lin, "Design and Implementation of Voting System Based on Ethereum," 2021 International Conference on Intelligent Computing, Automation and Applications (ICAA), Nanjing, China, 2021, pp. 610-617, doi: 10.1109/ICAA53760.2021.00112

[3] Kristof Lommers et al., “Confronting Machine Learning With Financial Research”.2022, DOI: 10.3905/jfds.2021.1.068

[4] Lukas Ryll and Sebastian Seidens, “Evaluating the Performance of Machine Learning Algorithms in Financial Market Forecasting: A Comprehensive Survey”, https://arxiv.org/abs/1906.07786 ,2019 https://doi.org/10.48550/arXiv.1906.07786

[5] Sonkavde, G., Dharrao, D. S., Bongale, A. M., Deokate, S. T., Doreswamy, D., & Bhat, S. K. (2023). Forecasting Stock Market Prices Using Machine Learning and Deep Learning Models: A Systematic Review, Performance Analysis and Discussion of Implications. International Journal of Financial Studies, 11(3), 94. https://doi.org/10.3390/ijfs11030094

[6] A. Saini, N. Singh, R. K. Singh and M. K. Sachan, "Financial Time Series Prediction on Apple Stocks Using Machine and Deep Learning Models," 2024 International Conference on Electrical, Computer and Energy Technologies (ICECET, Sydney, Australia, 2024, pp. 1-6, doi: 10.1109/ICECET61485.2024.10698514

[7] Ayitey Junior, M., Appiahene, P., Appiah, O. et al. “Forex market forecasting using machine learning: Systematic Literature Review and meta-analysis”. J Big Data 10, 9 (2023). https://doi.org/10.1186/s40537-022-00676-2

[8] Dr. Gurjeet Singh ,”Machine Learning Models in Stock MarketPrediction”. https://arxiv.org/abs/2202.09359,2022, https://doi.org/10.35940/ijitee.C9733.0111322

[9] Jyotirmayee Behera et al., An approach to portfolio optimization with time series forecasting algorithms and machine learning technique”.2024, https://doi.org/10.1016/j.asoc.2025.112741

[10] Kazuki Amagai and Tomoya Suzuki,’ Long-Term Modeling of Financial Machine Learning for Active Portfolio Management”, https://arxiv.org/abs/2301.12346, 2023, https://doi.org/10.48550/arXiv.2301.12346

[11] Markowitz, H. Portfolio Selection. The Journal of Finance, 7, 77-91. (1952) https://doi.org/10.1111/j.1540-6261.1952.tb01525.x

[12] A. Saini, N. Singh, R. K. Singh and M. K. Sachan, "Financial Time Series Prediction on Apple Stocks Using Machine and Deep Learning Models," 2024 International Conference on Electrical, Computer and Energy Technologies (ICECET, Sydney, Australia, 2024, pp. 1-6, doi: 10.1109/ICECET61485.2024.10698514

[13] S. Kumar, N. Pal and A. M. Tripathi, "Improving Long Short-Term Memory (LSTM)-Based Stock Market Price Predictions in the Machine Learning Era," 2024 IEEE International Conference on Computing, Power and Communication Technologies (IC2PCT), Greater Noida, India, 2024, pp. 923-928, doi: 10.1109/IC2PCT60090.2024.10486391

[14] Anshuman Behera et al., “Stock Price Prediction using Machine Learning”.2022, https://doi.org/10.1109/MLCSS57186.2022.00009