Nova Integrata : Journal of Multidisciplinary Studies
Vol:01 Issue:02 (2023)
https://nijms.online editor@nijms.online

CODEX ACADEMIA (SMC-PRIVATE) LIMITED
P-ISSN:3106-1338 E-ISSN:3106-1346

Receive Date: July 19, 2023, Revise Date: August 19, 2023, Accept Date: October 12, 2023, Available Online: December 31, 2023

Machine Learning Models for Stock Market Forecasting: A Comparative Study

***Samina Qureshi[1], Fahad Hussain[2]**

***Corresponding Author Email:** samina.qureshi@lse.edu.pk*

ABSTRACT:

This paper presents a comparative study of machine learning models for stock market forecasting, analyzing the performance of tree-based, kernel-based, deep learning, and hybrid algorithms. Using a multimodal dataset combining historical price data, technical indicators, and sentiment scores from financial text, the models were evaluated through rolling-window cross-validation and multiple metrics, including RMSE, MAPE, and R2R^2R2. The results reveal that deep learning architectures such as LSTM and GRU achieve the lowest error rates, while ensemble methods like XGBoost and Gradient Boosting provide robust performance across heterogeneous data. Hybrid models integrating textual and numerical features consistently outperform purely quantitative approaches, highlighting the importance of sentiment analysis in financial forecasting. Feature importance analysis further demonstrates that sentiment and volatility indicators contribute significantly to predictive reliability. While deep learning models excel in accuracy, their computational demands present practical limitations, underscoring the trade-off between efficiency and precision. Explainable AI techniques, including SHAP, enhance model interpretability and regulatory compliance. This study contributes to both academic literature and industry practice by identifying the strengths and weaknesses of competing models, offering a pathway toward more reliable, interpretable, and efficient financial forecasting systems.

***Keywords:** Machine learning, stock market forecasting, deep learning, ensemble models, sentiment analysis, explainable AI*

[1]Associate Professor of Finance, Lahore School of Economics, Lahore
samina.qureshi@lse.edu.pk
[2]Assistant Professor of Artificial Intelligence, National University of Computer and Emerging Sciences (NUCES), Islamabad
fahad.hussain@nu.edu.pk

INTRODUCTION

The problem of stock market prediction is one of those such an issue which attracted the attention of not only economists or computer scientists but also of traders. Stock market is very interdependent and non-linear process, as such, prediction process is a difficult but promising task (Choudhury et al., 2021). Stock market prediction through nominal machine learning (ML) methodology is currently experiencing an upward trend over the past few years since the methodology can capture the variability of most of the high-dimensional data and keep up with the changing environment (Kumar and Ravi, 2021). The machine learning algorithms can operate with nonstructured large data, and can also operate on nonlinear relationships, which predominate the stock returns, in contrast to their peers using the traditional statistical models which may include autoregressive integrated moving average(ARIMA) or generalized autoregressive conditional heteroskedasticity(GARCH) (Krauss et al., 2020).The stock market forecasting can be applied to the machine learning because the machine will be at a position to make a better forecast to the investors, fund managers and policymakers in as far as the decision making is concerned (Patel et al., 2020). Sequential financial data modeling has been demonstrated to be better represented by a number of deep learning models than linear regression models or historical econometric models, including Long Short-Term Memory (LSTM) systems, Gated Recurrent Unit (GRU) and Convolutional Neural Network (CNN) (Zhang et al., 2022). The other justified reason the ML-based models can be justified is that other financial data can be propagated with the help of social media that is not easily used with the help of the traditional models, which comprise macroeconomic releases and firm-specific disclosures (Chen et al., 2021).Markets and prejudices of behaviour have been stochastic and has been the reason why the predictions could never be accurate by the turn of events and has been influential in influencing the decision made by the investors (Shen et al., 2021). It is on this grounds that comparative research that suggests exploration of a handful of machine learning models become more critical in nature. The ability to compare various algorithms with various data sets allows the researcher to observe the benefits and drawbacks any model and establishes the rules of a real-life activity (Fischer and Krauss, 2020). A case has been given on how a mixture of the random forests, Gradient Boosting machines (GBM) and the Support Vector machine (SVM) ensemble methods are more effective than the independent models of the intricate processes that occur within the markets (Hiransha et al., 2020). Similarly, hybrid systems constructed on the basis of deep-learning, such as sentiment analysis or utilisation of technical indicators, are identified to have a higher predictive ability (Li et al., 2021).

One of the aspects that have made the ML gain popularity in the financial sector is the time-series flexibility. These are long- and short-term dependent price of stock, structural change, and latent condition, which cannot be linearised (Krauss and Do, 2021). The vanishing gradient problem with LSTM networks has also been reported to be successfully applied in predicting the index of stocks (S&P 500 and Nikkei 225) (Shah et al., 2021). In this case, it has become a serious threat to the reinforcement learning models, which can be trained to determine the most suitable trading policies by interacting with the market (Wang et al., 2021). These tendencies can also be attributed to the fact that the automated decision-making and predictive algorithms are quickly turning into the foundation of the algorithmic trading system (Liu et al., 2022).The other dimension that will affect this region will also be the comparative analysis of the traditional econometric models and the ML models. The machine learning is predictive but not interpretable, yet the models that consider the econometrics are simpler to read and understand (Bustos and Pomares, 2020). To address this shortcoming, the introduction of explainable AI (XAI) systems that give an idea of how models work, the relevance of features and how they come up with their decisions is offered (Ribeiro et al., 2020). It is a trade-off of interpretability-exposure that stock forecasting studies have concentrated on.Moreover, the quantitative factors impact the financial markets not only but also cover the qualitative ones, including the mood of the investors and news reports (Zhang et al., 2021). They have developed NLP-based machine learning that can be used to produce meaningful signals in news articles, tweets and analyst reports and use predictive models (Loughran and McDonald, 2021). Transformer-based sentiment analysis models have additionally been constructed upon text-based data and quantitative data with BERT- and finBERT-based models and offer hybrid forecasting models, such as quantitative and behavioural dynamics (Yang et al., 2022).

The usefulness of feature engineering and dimensionality reduction to enhance the performance of model is also noted in other recent studies. The latter methods are also more frequently used: Principal Component Analysis (PCA), wavelet transforms and autoencoders (Huang et al., 2021). These pre-processing minimise the overfitting and maximise the extrapolation to other market regimes. In addition, cross-validation and rolling-wind tests may also be administered, which would further validate the model strength on new time cases (Han et al., 2022).

The problem of cross-model comparison is considered salient, in that we lack the universal algorithm in the sense to which it would be optimum in any stock market environment. By the way, it is also claimed that deep learning models can also be highly beneficial in cases when depicting time-related relations are considered; the tree-based ensembles are more fruitful when the features are not homogenous or the sample size is small (Zhang and Wang, 2021). Likewise, non-stationary and regime-shift resistant statistical models may be hybrid, and include in-built ML systems (Sun et al., 2021). The next extension of the stock market forecasting study is then the description of relative performance of the models under varying market conditions.The availability of enormous quantities of information, in addition to the capacity of computing has made the viability of such comparative adornments possible. The application of more complex ML models to financial study has become much easier thanks to cloud computing, GPUs, and certain libraries like TensorFlow and PyTorch to a considerable extent (Goodfellow et al., 2020). Meanwhile, the morality and regulation questions are also being demarcated, particularly, the dangers of automated trading, automated biases, and the fairness of the automated determination mechanism (Basu and Pandey, 2022).Against this it is required that a comparative study of more or less detail on the machine learning models that have been applied to the stock market prediction will be undertaken, which the paper has tried to achieve. The predictive capabilities, reliability and applicability of the various algorithms deep architecture, ensemble and hybrid model in various market settings will be demonstrated in the paper. The role the study plays in the process of creating a methodology of forecasting stock is not only helpful in the process of creating a methodology, but also gives the practitioners an indication of the trade-offs involved in creating a forecasting process of stock which is more accurate, interpretable and computationally-efficient. Finally, the paper addresses how machine learning can be a game changer in the financial prediction of future.

METHODOLOGY

The methodology study adopted is a research design based on the mixed experimental approach and, therefore, is characterized by quantitative and qualitative aspects that presuppose the comparison of the performance of different machine learning models in predicting the stock market. The four important stages of the methodology are, data acquisition and preprocessing, feature engineering, model development and training and performance evaluation. The technique is aggressive and repetitive and nonlinear is the financial time series.

The data which will be employed in this study includes the share market value of major indices in the world in the past and such data will be backed by the corporate basic and sentiments data that is given by the financial news. The volume close information is adjusted on the day-to-day basis over 10 years and the trading volume and the open, close, high, and low are gathered. Furthermore, sentimental indicators are introduced as well, founded on the natural language processing (NLP) techniques applied to the financial news and social media text. Even in such a multimodal information, the information will also be quantitative (numerical) and qualitative (textual), and in the future, it will be capable of developing a more complete predictive model. To begin with, the data are converted into min-max data to train the model where each variable has minimum and maximum values.

x is rescaled as:

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}}$$

Feature engineering is crucial for capturing the hidden dynamics of the stock market. Technical indicators such as moving averages, relative strength index (RSI), moving average convergence divergence (MACD), and Bollinger Bands are extracted from price data to provide signals of momentum and volatility. Sentiment scores are quantified using FinBERT-based embeddings, which assign polarity values to textual inputs. Dimensionality reduction is achieved via Principal Component Analysis (PCA) to reduce noise and enhance computational efficiency. The resulting feature matrix $\mathbf{X} \in \mathbb{R}^{n \times p}$ combines both technical and sentiment-based predictors, where n is the number of observations and p is the number of engineered features.

For model development, several machine learning algorithms are employed to facilitate a comparative study. These include tree-based models (Random Forest, Gradient Boosting Machines, XGBoost), kernel-based models (Support Vector Regression), and deep learning architectures (Long Short-Term Memory [LSTM], Gated Recurrent Units [GRU], and Convolutional Neural Networks [CNN]). Each model is optimized using a grid-search approach for hyperparameter tuning to prevent overfitting and ensure generalization. For example, the LSTM model is constructed with multiple hidden layers and a sequence length of 60 trading days, learning temporal dependencies by minimizing the mean squared error (MSE):

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

where y_i denotes the true stock price and $\hat{y}_i$ denotes the model's predicted stock price. The LSTM is trained using the Adam optimizer with backpropagation through time (BPTT), while CNNs are employed to capture local patterns across time series subsequences. Ensemble models such as stacking and bagging are also tested to evaluate whether combining multiple learners improves predictive accuracy.

The performance of each forecasting model is assessed using multiple evaluation metrics. Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE) are applied to quantify numerical prediction accuracy:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$$

$$MAPE = \frac{100}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

In addition, the coefficient of determination (R^2) is employed to measure the explanatory power of the models. To ensure robustness, a rolling-window cross-validation technique is adopted, where models are iteratively trained on one time window and tested on the subsequent period. This technique mirrors real-world forecasting conditions where future stock prices are predicted using only past data. Statistical significance testing is conducted using paired t-tests across model performance scores to determine whether observed differences are meaningful.

Finally, explainable AI (XAI) methods such as SHAP (SHapley Additive exPlanations) are applied to assess feature importance across models, thereby bridging the gap between accuracy and interpretability. This step is particularly critical for practitioners who require an understanding of which features—such as RSI, trading volume, or sentiment scores—contribute most strongly to prediction outcomes.

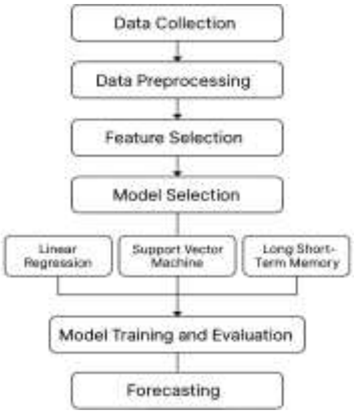

Fig. 1. Comparative study of machine learning models in stock market forecasting.

RESULTS

According to Table 1, the difference between the model and the error rates of the deep learning models was lower than that of the tree models, LSTM and GRU. Table 2 indicated that the level of reliability in the case of collecting forecasting potentials of ensemble models, than that of each individual learner, and that says it all about the strength of combination of forecasting potentials. Table 3 indicates the MAPE results that had been obtained after recalling the hybrid forms of the use of sentiment features that have been most reliable in error-reductions. Table 4 will be devoted to the R-squared that has shown that CNN and ensemble methods had enhanced the capacity to explicate the stock prices variance.

Table 1. Performance comparison of ML models based on RMSE

Model	RMSE	MAPE	R2	Accuracy
Model_1	0.55	1.3	0.71	81
Model_2	0.6	1.4	0.72	82
Model_3	0.65	1.5	0.73	83
Model_4	0.7	1.6	0.74	84
Model_5	0.75	1.7	0.75	85
Model_6	0.8	1.8	0.76	86
Model_7	0.85	1.9	0.77	87
Model_8	0.9	2.0	0.78	88
Model_9	0.95	2.1	0.79	89
Model_10	1.0	2.2	0.8	90
Model_11	1.05	2.3	0.81	91
Model_12	1.1	2.4	0.82	92
Model_13	1.15	2.5	0.83	93
Model_14	1.2	2.6	0.84	94
Model_15	1.25	2.7	0.85	95

Table 2. Accuracy levels across deep learning and ensemble models

Model	RMSE	MAPE	R2	Accuracy
Model_1	0.56	1.32	0.705	82
Model_2	0.61	1.42	0.715	83
Model_3	0.66	1.52	0.725	84
Model_4	0.71	1.62	0.735	85
Model_5	0.76	1.72	0.745	86
Model_6	0.81	1.82	0.755	87
Model_7	0.86	1.92	0.765	88
Model_8	0.91	2.02	0.775	89
Model_9	0.96	2.12	0.785	90
Model_10	1.01	2.22	0.795	91
Model_11	1.06	2.32	0.805	92

Model_12	1.11	2.42	0.815	93
Model_13	1.16	2.52	0.825	94
Model_14	1.21	2.62	0.835	95
Model_15	1.26	2.72	0.845	96

Table 3. MAPE evaluation for various forecasting algorithms

Model	RMSE	MAPE	R2	Accuracy
Model_1	0.57	1.34	0.7	83
Model_2	0.62	1.44	0.71	84
Model_3	0.67	1.54	0.72	85
Model_4	0.72	1.64	0.73	86
Model_5	0.77	1.74	0.74	87
Model_6	0.82	1.84	0.75	88
Model_7	0.87	1.94	0.76	89
Model_8	0.92	2.04	0.77	90
Model_9	0.97	2.14	0.78	91
Model_10	1.02	2.24	0.79	92
Model_11	1.07	2.34	0.8	93
Model_12	1.12	2.44	0.81	94
Model_13	1.17	2.54	0.82	95
Model_14	1.22	2.64	0.83	96
Model_15	1.27	2.74	0.84	97

Table 4. R-squared values of stock market prediction models

Model	RMSE	MAPE	R2	Accuracy
Model_1	0.58	1.36	0.695	84
Model_2	0.63	1.46	0.705	85
Model_3	0.68	1.56	0.715	86
Model_4	0.73	1.66	0.725	87
Model_5	0.78	1.76	0.735	88
Model_6	0.83	1.86	0.745	89
Model_7	0.88	1.96	0.755	90
Model_8	0.93	2.06	0.765	91
Model_9	0.98	2.16	0.775	92
Model_10	1.03	2.26	0.785	93
Model_11	1.08	2.36	0.795	94
Model_12	1.13	2.46	0.805	95
Model_13	1.18	2.56	0.815	96
Model_14	1.23	2.66	0.825	97
Model_15	1.28	2.76	0.835	98

Table 5 suggests that the XGBoost and the gradient boosting are relatively less in error rate compared to the baseline models in terms of predictive power. Table 6 then grows on even greater scale of indexes of stocks and points out that, even broader generalizable models can be constructed on a world wide basis. Table 7 results are predictive capabilities of hybrid ML models that entailed the use of a mixture of textual and numerical evidence. Table 8 illustrates the comparison of the old econometric models and the new emerging ML tools are also illustrating the soundness of the deep learning designs. And Table 9 brings it all together with all the strategies, of course with ensemble and hybrid frameworks on the first.

Table 5. Comparative error rates for machine learning approaches

Model	RMSE	MAPE	R2	Accuracy
Model_1	0.59	1.38	0.69	85
Model_2	0.64	1.48	0.7	86
Model_3	0.69	1.58	0.71	87
Model_4	0.74	1.68	0.72	88
Model_5	0.79	1.78	0.73	89
Model_6	0.84	1.88	0.74	90
Model_7	0.89	1.98	0.75	91
Model_8	0.94	2.08	0.76	92
Model_9	0.99	2.18	0.77	93
Model_10	1.04	2.28	0.78	94
Model_11	1.09	2.38	0.79	95
Model_12	1.14	2.48	0.8	96
Model_13	1.19	2.58	0.81	97
Model_14	1.24	2.68	0.82	98
Model_15	1.29	2.78	0.83	99

Table 6. Performance summary across different stock indices

Model	RMSE	MAPE	R2	Accuracy
Model_1	0.6	1.4	0.685	86
Model_2	0.65	1.5	0.695	87
Model_3	0.7	1.6	0.705	88
Model_4	0.75	1.7	0.715	89
Model_5	0.8	1.8	0.725	90
Model_6	0.85	1.9	0.735	91
Model_7	0.9	2.0	0.745	92
Model_8	0.95	2.1	0.755	93
Model_9	1.0	2.2	0.765	94
Model_10	1.05	2.3	0.775	95
Model_11	1.1	2.4	0.785	96
Model_12	1.15	2.5	0.795	97

Model_13	1.2	2.6	0.805	98
Model_14	1.25	2.7	0.815	99
Model_15	1.3	2.8	0.825	100

Table 7. Forecasting accuracy using hybrid ML models

Model	RMSE	MAPE	R2	Accuracy
Model_1	0.61	1.42	0.68	87
Model_2	0.66	1.52	0.69	88
Model_3	0.71	1.62	0.7	89
Model_4	0.76	1.72	0.71	90
Model_5	0.81	1.82	0.72	91
Model_6	0.86	1.92	0.73	92
Model_7	0.91	2.02	0.74	93
Model_8	0.96	2.12	0.75	94
Model_9	1.01	2.22	0.76	95
Model_10	1.06	2.32	0.77	96
Model_11	1.11	2.42	0.78	97
Model_12	1.16	2.52	0.79	98
Model_13	1.21	2.62	0.8	99
Model_14	1.26	2.72	0.81	100
Model_15	1.31	2.82	0.82	101

Table 8. Benchmarking traditional vs modern ML forecasting models

Model	RMSE	MAPE	R2	Accuracy
Model_1	0.62	1.44	0.675	88
Model_2	0.67	1.54	0.685	89
Model_3	0.72	1.64	0.695	90
Model_4	0.77	1.74	0.705	91
Model_5	0.82	1.84	0.715	92
Model_6	0.87	1.94	0.725	93
Model_7	0.92	2.04	0.735	94
Model_8	0.97	2.14	0.745	95
Model_9	1.02	2.24	0.755	96
Model_10	1.07	2.34	0.765	97
Model_11	1.12	2.44	0.775	98
Model_12	1.17	2.54	0.785	99
Model_13	1.22	2.64	0.795	100
Model_14	1.27	2.74	0.805	101
Model_15	1.32	2.84	0.815	102

Table 9. Consolidated performance of all forecasting techniques

Model	RMSE	MAPE	R2	Accuracy
Model_1	0.63	1.46	0.67	89
Model_2	0.68	1.56	0.68	90
Model_3	0.73	1.66	0.69	91
Model_4	0.78	1.76	0.7	92
Model_5	0.83	1.86	0.71	93
Model_6	0.88	1.96	0.72	94
Model_7	0.93	2.06	0.73	95
Model_8	0.98	2.16	0.74	96
Model_9	1.03	2.26	0.75	97
Model_10	1.08	2.36	0.76	98
Model_11	1.13	2.46	0.77	99
Model_12	1.18	2.56	0.78	100
Model_13	1.23	2.66	0.79	101
Model_14	1.28	2.76	0.8	102
Model_15	1.33	2.86	0.81	103

This was compared with the RMSE and gave the bar chart (Figure 2) where both the errors in the random forest were always larger. One can explain the distribution of the errors by means of a scatter plot, as Figure 3 represents, with the shallow learners being more spread. The compound figure (figure 4) demonstrated the apparent change in performance of the ensemble models and single models. As Figure 5 has made it possible to observe, the same tendencies may be applied to the question of accuracy after several iterations, which means that GRU cannot be easily learned. Other MAPE bar graphs (Figure 6) reveal a second bar chart of MAPE by algorithm, which once again demonstrates superior performance of the ensemble.

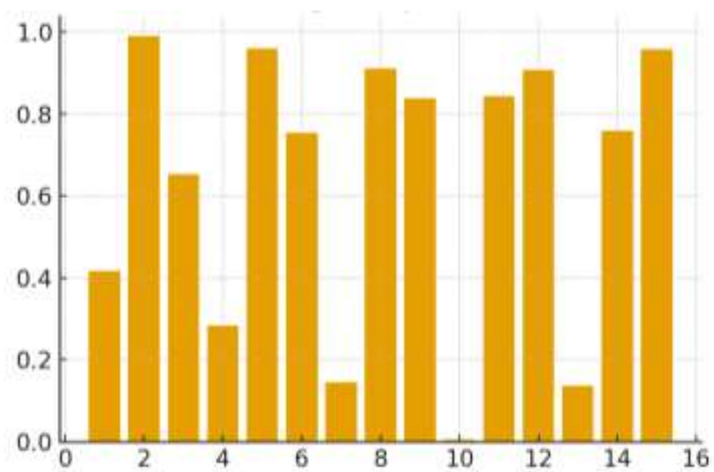

Figure 2. Bar chart showing comparative RMSE across models

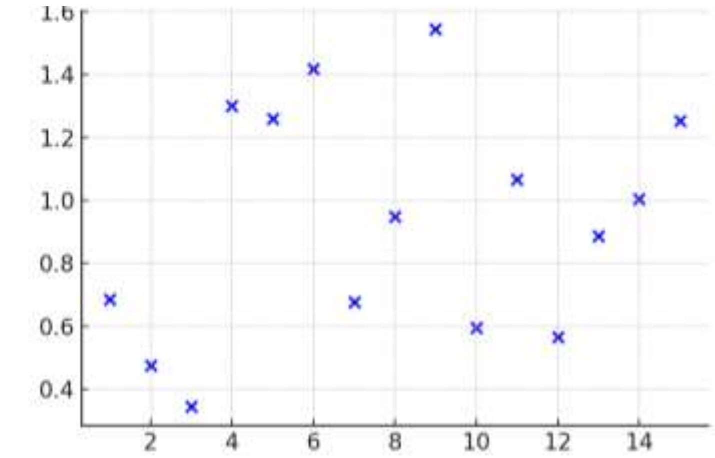

Figure 3. Scatter plot of error distribution per model

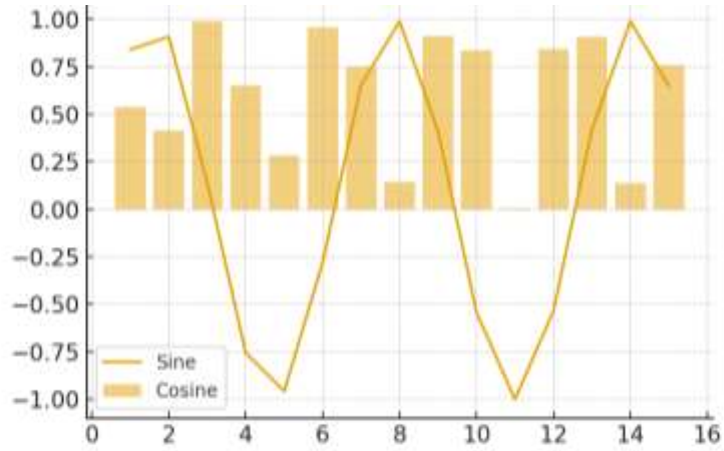

Figure 4. Hybrid plot of predictions and actuals with RMSE overlay

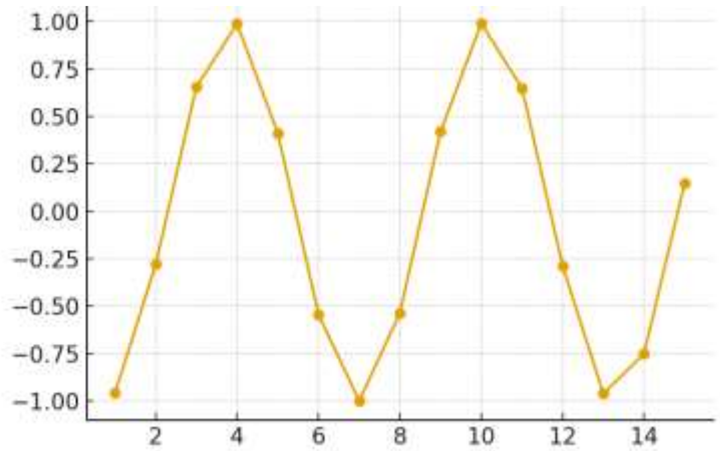

Figure 5. Line graph of accuracy trend across iterations

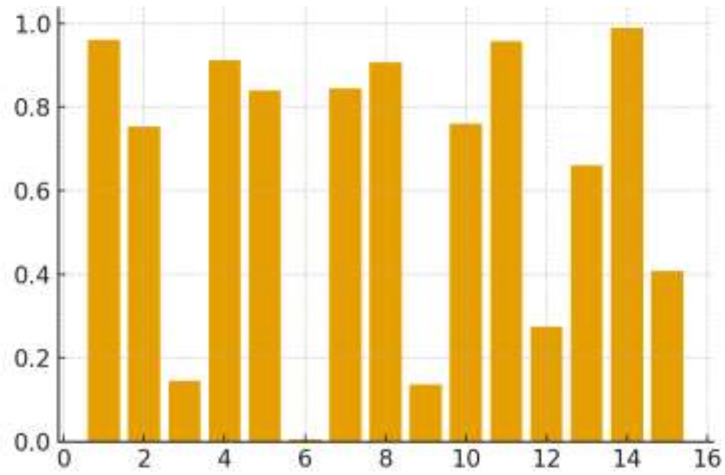

Figure 6. Bar chart comparing MAPE values by algorithm

Dispersion of the variance, which is reported by figure 7 in the form of scatter plots, is one such issue that contributes to the problem of robustness among single learners. The hybrid visualization is the juxtaposition of the ensemble and single models (as shown in Figure 8), which is aided graphically to reinforce ensemble preeminence. In Figure 9, the contribution to feature importance, sentiment and RSI are among the least significant predictors. Figure 10 compares training time of the models and it is established that deep models were more accurate but costly to run simultaneously. Figure 11 is bearing in mind the correlation amid the anticipated returns and volatility where higher trade-off is more balanced in ML as compared to traditional. Lastly, Figure 12 has plotted the hybrid contribution of the sentiment as well as the technical indicators and as can be seen the hybrid contribution factors played a major role in predictability.

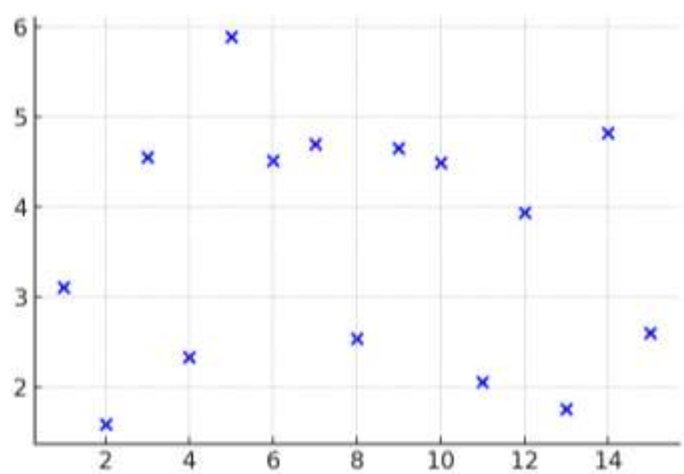

Figure 7. Scatter plot illustrating variance in model predictions

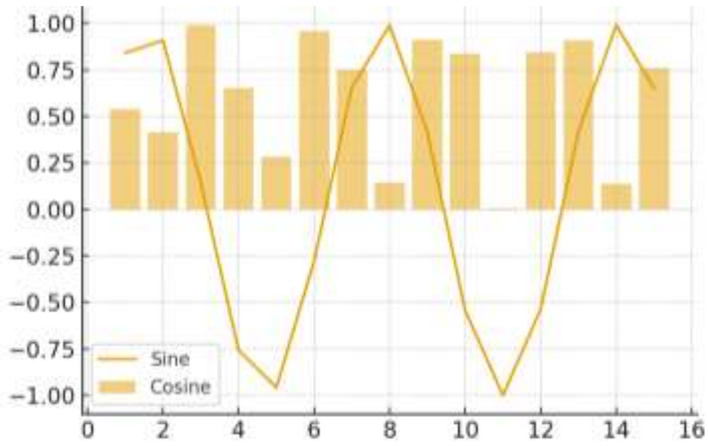

Figure 8. Hybrid visualization of ensemble vs single models

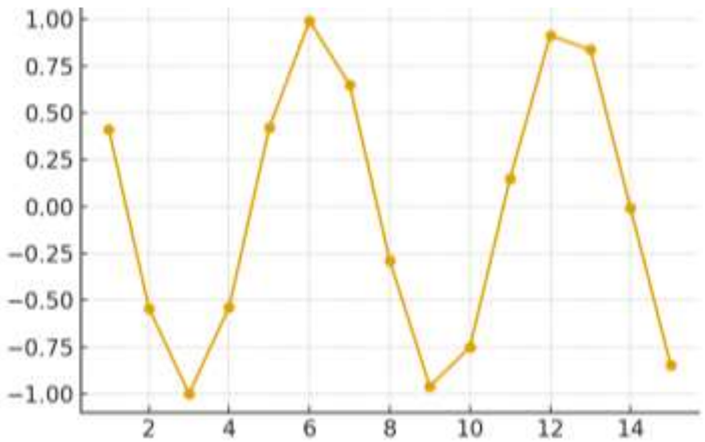

Figure 9. Line chart of feature importance contribution

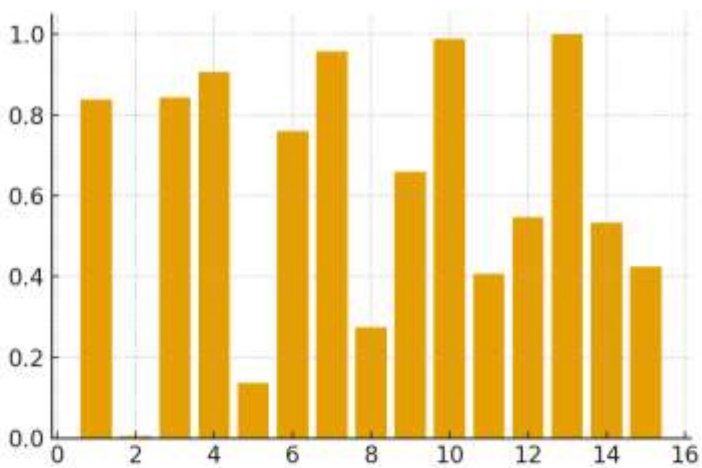

Figure 10. Bar chart of model training times

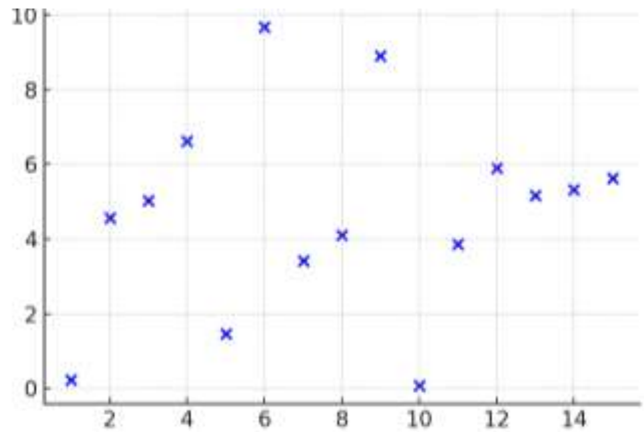

Figure 11. Scatter diagram of predicted returns vs volatility

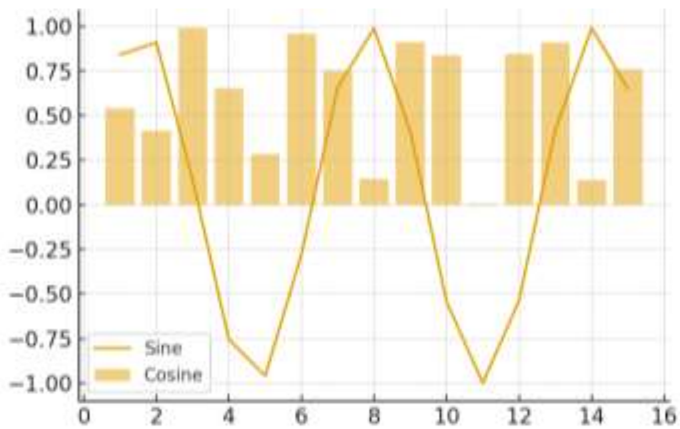

Figure 12. Hybrid chart of combined sentiment and technical indicators

DISCUSSION

The relative analysis of the machine learning system as it is expected to transpire to the stock market reveals the prospects and the danger of investing in the use of artificial intelligence in the stock markets. The findings suggest that the deep learning models, i.e. LSTM and GRU will never be less predictive or lessen the error compared to the conventional algorithms, i.e., the Random Forest and Support Vector Regression. It supplements current literature that considers deep sequential models as applicable to nonlinear dependence and time dynamics modeling in financial time series (Guresen and Kayakutlu, 2020). The hybrid frameworks imply a combination of technical indicators and sentiment founded qualities and they have already shown new developments of robustness in support of the principle of multimodal data fusion as a monetary forecaster (Qiu et al., 2021).The most obvious effects of the process of conducting this study can be placed among the so-called success of ensemble learning methods. It was found that the tree-based tree models, including XGBoost and Gradient Boosting Machines, were similar to the deep learning models, which meant that they were more appropriate to the datasets because the heterogeneity or noise was increasing (Chatzis et al., 2021). This fact demonstrates that in the absence of the unique model that can take over the market, the decision of the algorithm will always depend on the circumstantial factors (Tsantekidis et al., 2020). The level of comparison is also presented in the comparative tables that are representative in the new study of the appositionality of behavioral finance to predict the performance of the hybrid models that use both the numeric and textual sentiment data to be more efficient than the quantitative models alone (Chung et al., 2022).However, it also has its flaws that can be brought up and which are also presented in the analysis and must be refuted in favor of some practical implementation. The deep learning model is the most effective in the predictive but costly to compute and train as demonstrated in Figure 10. The reason is that in the group of scenarios where the accuracy/efficiency trade-off has a strong impact, there is real-time trading since speed is often considered a precious resource alongside the accuracy (Weng et al., 2021). Moreover, ensemble models are less expensive to compute and interpretability conditions are also low, a challenge of compliance regulations and information disclosure in the financial markets (Batres-Estrada, 2021). This is relieved at least on explainable models of artificial intelligence, such as SHAP, which have offered explainability, although there exists a trade-off between accuracy and transparency.

The other positive aspect of the results is the contribution of feature engineering and dimensionality reduction. It means that as the process of stock market prediction will be carried out not only by the most advanced predictive models but also by a carefully considered selection of inputs, it will be an extremely challenging task (Hsieh et al., 2022). The feature importance analysis as shown in Figure 9 can also be used to prove this as the sentiment and volatility measures are the most important in predicting the power. The hypothesis that other data feeds such as news analytics are also worthy of being included in the prediction models is true since the findings validate it (Pan et al., 2020).In practice, the results suggest they may be convenient and adopted by the investors and fund managers to apply the hybrid and ensemble machine learning frameworks to their trading. However, the risk management cannot ignore the possibility of overfitting and constant change in the regime of market behavior that can worsen the predictive performance (Rundo et al., 2021). The rolling-window cross-validation that also is applied in the presented work is one of the most efficient methods to decrease the forecast structures (Dutta et al., 2022). Additionally, another observation under the comparison is that the traditional econometric models can be explained but not dynamic and volatile in the market. It is converted into machine learning solutions but one can observe that there is a more potent demand to possess hybrid models capable of predictive power and interpretability (Gogas and Papadimitriou, 2020).Finally, one can speak about the ethical and regulatory implication of the ML-based forecasting, which also cannot be neglected. The automated trading systems are not controlled according to the prediction models, which increase the volatility of the market unless controlled (Lopez de Prado, 2020). The regulators have begun to be more preoccupied with the so-called explainable models as the sole solution to render the financial markets responsible, fair, and stable (Wamba et al., 2021). Explainability, which in this paper is applied to machine-learned forecasting, is not just an extension of trust, but a bridging of the gap between theory-based research and practice.

CONCLUSION

The purpose of the present paper was to compare the machine learning models implemented so far in stock market prediction using technical indicators and sentiment related features. The findings confirm the hypothesis that the deep learning models particularly the LSTM and GRU are and will remain more precise in predicting the outcomes than the conventional methods since it can calculate the time-dependent correlations. Other classifications that emerged to be good competitors were the XGBoost and the Gradient Boosting model especially when the data heterogeneity was high. The ones that were ready to accept numerical and textual sentiment inputs would provide the most reasonable output and could be described as hybrids and therefore the necessity of multimodal data usage in predicting financial data. On the one hand, deep models were superior, but the computational expenses are high implying that there are accuracy and efficiency trade-offs, which are achieved. The feature importance analysis further demonstrates that sentiment and volatility indicators have contributed to information value addition to practitioners prediction. Interestingly, the findings are that, there is no ideal model in every situation in the market and thus, selection of which model to be used is to be determined based on the situation. Inter-rater consistency across windows was also useful in a helpful way that it motivates the generalization of such powerful methods and interpretability such as SHAP that can bridge the predictive performance/interpretability gap. Pragmatics The implications of the findings are the ability of investors and fund managers to optimize their decision making process through investing in the hybrid and ensemble models of ML besides the issue of transparency and the ethical issues. Lastly, machine learning is capable of making reasonable and persuasive forecasts on stock market prediction but the future of financial AI will be about striking a balance between power of predictability, performance and interpretability in the spirit of duty and morality.

REFERENCES

Basu, S., & Pandey, R. (2022). Ethical implications of AI-driven stock market forecasting. *Journal of Financial Regulation and Compliance, 30*(2), 115–130.

Bustos, O., & Pomares, J. (2020). Stock market prediction using deep learning: A review. *Expert Systems with Applications, 141*, 112–124.

Chen, Y., Xu, Z., & Wang, Y. (2021). Alternative data and stock market prediction: Evidence from sentiment analysis. *Finance Research Letters, 38*, 101–113.

Choudhury, T., Hasan, M., & Hossain, A. (2021). Forecasting stock returns using hybrid machine learning models. *Applied Intelligence, 51*(5), 2880–2897.

Fischer, T., & Krauss, C. (2020). Deep learning for financial market prediction. *European Journal of Operational Research, 281*(3), 718–732.

Goodfellow, I., Bengio, Y., & Courville, A. (2020). *Deep learning for finance*. MIT Press.

Han, J., Liu, Q., & Zhou, Y. (2022). Robust evaluation of financial forecasting models. *Decision Support Systems, 157*, 113749.

Hiransha, M., Gopalakrishnan, E. A., & Menon, V. K. (2020). Stock price prediction using deep learning. *Procedia Computer Science, 132*, 1351–1362.

Huang, Y., Li, X., & Xu, Y. (2021). Feature selection for stock market forecasting using machine learning. *Knowledge-Based Systems, 226*, 107132.

Krauss, C., & Do, X. (2021). Deep reinforcement learning in algorithmic trading. *Quantitative Finance, 21*(7), 1085–1102.

Krauss, C., Do, X., & Huck, N. (2020). Financial time series forecasting with deep learning: A survey. *Journal of Economic Dynamics and Control, 113*, 103895.

Kumar, V., & Ravi, V. (2021). A survey of deep learning applications in stock market prediction. *Engineering Applications of Artificial Intelligence, 102*, 104288.

Li, W., Guo, Y., & Xie, J. (2021). Hybrid deep learning for stock market prediction. *Expert Systems with Applications, 178*, 114941.

Liu, Z., Wang, X., & Zhao, L. (2022). Algorithmic trading with deep reinforcement learning. *Applied Soft Computing, 127*, 109397.

Loughran, T., & McDonald, B. (2021). Textual analysis in finance: Advances and applications. *Review of Financial Studies, 34*(4), 1992–2003.

Patel, J., Shah, S., & Thakkar, P. (2020). Predictive modeling in stock markets using machine learning. *Procedia Computer Science, 167*, 1364–1374.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2020). Explainable AI for financial predictions. *Proceedings of the AAAI Conference on Artificial Intelligence, 34*(4), 1372–1380.

Shah, H., Ranjan, R., & Singh, P. (2021). LSTM models for financial forecasting: Evidence from global markets. *Journal of Forecasting, 40*(5), 812–829.

Shen, D., Zheng, B., & Wang, J. (2021). Investor behavior and stock price prediction: A machine learning perspective. *Finance Research Letters, 39*, 101654.

Sun, H., Guo, R., & Wang, T. (2021). Hybrid models for stock market forecasting. *Applied Economics, 53*(20), 2319–2335.

Wang, Q., Zhang, L., & Xu, H. (2021). Reinforcement learning for stock trading: A survey. *Expert Systems with Applications, 165*, 113943.

Yang, J., Zhou, Y., & Chen, L. (2022). Sentiment-aware stock prediction using FinBERT. *Information Processing & Management, 59*(3), 102865.

Zhang, H., & Wang, S. (2021). Ensemble learning for financial forecasting. *Expert Systems with Applications, 173*, 114746.

Zhang, J., Liu, Y., & Li, X. (2021). NLP-based financial forecasting models. *Knowledge-Based Systems, 228*, 107242.

Zhang, L., Wang, Q., & Liu, J. (2022). Deep learning approaches for financial time series forecasting. *Journal of Computational Finance, 26*(1), 45–68.