Effectiveness of Machine Learning Algorithms in Predicting Financial Market Movements

***Abstract*— The challenging yet traditional field of stock market forecasting has garnered significant interest from data scientists and economists. Given the high risks associated with trading, where investors can lose part or all of their initial capital, there is a growing need for advanced methods to support investment decisions. By leveraging the capability to predict future stock performance patterns, the return on investment for short-term trading can potentially be increased. This research reviews various machine learning techniques and their effectiveness with Apple stock data spanning from 2010 to 2021. Among several machine learning algorithms, the SVM algorithm is shown to be the most accurate, with an accuracy rate of 76%. However, the effectiveness of machine learning systems in predicting stock prices is influenced by several factors. The inherent volatility and unpredictability of financial markets make it challenging for models to anticipate sudden, unforeseen events. Overfitting is a common issue where models that perform well on historical data fail to generalize to new, unseen data because they capture noise instead of genuine market trends. Data quality issues, such as noisy or incomplete data, further complicate predictions. Additionally, the complexity of machine learning models can hinder interpretability, making it easier for analysts to trust and utilize their outcomes. This research underscores the necessity for careful and ongoing refinement of ML techniques in stock price forecasting.**

Keywords*— *Stock market, artificial intelligence, forecasting, machine learning, prediction, financial analysis.

I. Introduction

The stock market influences the nation's economy in two primary ways: by providing liquidity and facilitating price discovery. A strong stock market significantly benefits economic activity through the effective use of financial resources and luring in foreign investment. Traditional analytical methods, which focus solely on financial and economic perspectives, fall short of predicting future stock values. Social attitudes consistently affect the stock market, making future predictions complex due to the numerous variables involved [1]. Currently, machine learning (ML) is rapidly advancing and revolutionizing a wide array of applications. As ML continues to evolve, it becomes more crucial for driving innovation, enhancing productivity, and supporting more informed decision-making across various sectors [2]. ML is also transforming financial market research and prediction by utilizing extensive datasets and advanced algorithms to uncover patterns and trends that traditional methods frequently overlook. In the financial markets, ML is applied in risk management to assess market conditions and predict potential future events, thereby reducing losses. Additionally, it is used in algorithmic trading to execute transactions in real-time at optimal moments. Stock market price forecasting entails predicting future stock prices using a range of analytical approaches [3]. This process utilizes historical data, statistical models, and occasionally machine learning techniques to uncover patterns and trends that may suggest future price movements. Engaging in the trading and selling of cryptocurrencies or other financial assets carries substantial risks, including the potential loss of invested funds, and may not be suitable for all traders. The value of stock assets is highly volatile and can be influenced by external factors such as financial, political, or regulatory changes [4].

Before investing in financial instruments or cryptocurrencies, individuals should understand the associated costs and risks, corresponding risk compassion, expertise threshold, and investing goals with extreme caution, and seek professional advice if necessary. Techniques such as Mean-Variance, Bayesian Asset Allocation, and other statistical methods are used to analyze this data. However, to improve stock market predictions, a more advanced approach is needed—one that considers a greater number of factors and leverages the ability of computers to rapidly identify and calculate patterns [5]. The foundational assessing strategy involves a numerical investigation of stock prices, interest rates, inflation, financial statements of listed companies, and corporate policies. These parameters are used to forecast future stock trading. While macro-level stock analysis can be useful, its forecast accuracy is not always reliable [6]. Investing in the stock market requires a substantial amount of information. The most crucial aspect of stock market investment is selecting stocks and evaluating their trends, along with understanding the company's philosophy, goals, and mission—all of which require extensive data. Cutting-edge technology like ML algorithms can analyze this vast amount of information, aiding in the selection of quality stocks in some cases [7]. Alternatively, technical indicators can be used to predict future stock trends based on three assumptions: similar situations are likely to recur, stock prices move according to certain patterns, and stock prices reflect all available information [8]. Advancements in ML now allow us to assume financial markets by computing current statistical data. Machine learning involves computer systems or software that can forecast stock prices by learning from existing data, identifying patterns in historical data, and generating desired outcomes. Current research highlights the potential for developing more accurate stock market predictions by analyzing various forecasting methodologies that were used in earlier research.

II. Literature Review

ML, one of the subsets of Artificial Intelligence (AI), is driving noteworthy advancements in the business sector. Stock investors employ several algorithms to make trading decisions on assets with high speed and effectiveness. Nonetheless, there are numerous conventional ML, neural network (NN) and deep learning (DL) techniques for predicting stock prices. Identifying the most frequently used methods for estimating stock prices is one of the most critical issues in current stock market prediction [9]. These algorithms often become increasingly sophisticated as they leverage artificial intelligence (AI) to adapt to various trading patterns. Algorithmic trading is evolving towards more effective machine learning (ML), which can instantly analyze massive volumes of data from multiple sources. To accurately predict stock prices, a range of machine learning methods is employed, including Autoregressive Integrated Moving Average (ARIMA), Singular Value Decomposition (SVD), Long Short-Term Memory Networks (LSTM), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN). These techniques allow us to estimate the value of the stock for the next day [10].

Recent studies indicate that stock market analysis techniques fall into two main categories: mathematical and artificial intelligence-based. AI technology involves machine learning algorithms, while mathematical technology encompasses statistical tools [11]. Additionally, It has been noted that the majority of the chosen research areas focus on evaluating the accuracy of stock market predictions using several ML methods [12]. In a study, [13] evaluating the success rates of various approaches, ANN (Artificial Neural Network) and ARIMA algorithms were used to forecast the NYSE index. The findings indicate that both models effectively predicted the trend of index progress using a regular dataset that spans from 1988 to 2011. However, the ANN model outperformed the ARIMA model. In another study [14] study, random forest (RF), and support vector machine methods were used to forecast the movements of the DJIA and NDX indexes in the US. The results, based on a daily dataset from 2012 to 2018, showed that the RF algorithm outperformed the support vector machine (SVM) method for the DJIA index, while the NDX index predictions were more accurate using the SVM method. The authors [15] employed SVMs, KNN, decision trees, and RF algorithms to predict the movement directions of the S&P 500 and FTSE 100. The study utilized a dataset of 6042 daily recordings spanning from March 7, 1995, to August 28, 2018. The research concluded that the RF algorithm was the most effective for forecasting market movements in industrialized nations. Additionally, an early reported study [16] examined the forecast accuracy rates of ML algorithms on the NASDAQ, NYSE, NIKKEI, and FTSE indexes utilizing the information of each day from March 24, 2010, and March 24, 2020. The findings revealed that the RF method had the best overall performance. Upon reviewing similar research in the literature, it was noted that one or more ML algorithms are commonly employed to forecast the trend of stock index movements in industrialized nations.

III. FUNDAMENTALS OF STOCK PREDICTION

Machine learning algorithms utilize extensive and diverse datasets to forecast changes in financial markets, incorporating both financial and non-financial data types. Key datasets include historical price and volume data of stocks, commodities, and other financial instruments, which are crucial for trend analysis and pattern detection. Additionally, economic data such as interest rates, inflation rates, and employment statistics are essential as they impact market dynamics. Sentiment data, gathered from financial reports, social media, and news stories, provides insights into investor and public mood, significantly influencing market movements. Corporate financial statements, which detail a company's financial performance and health, are also crucial. These statements include cash flow reports, income statements, and balance sheets. Most datasets in stock market prediction come with labels. For instance, characteristics such as the Relative Strength Index (RSI) and Moving Average Convergence Divergence (MACD) are included, with the technical evaluation that commonly employs the stock's closing price as the target objective. This data, presented in a time-series format, consists of continuous values. Conversely, the target value in fundamental inspection is typically a signal for making buy or sell decisions. This method involves components like financial statements and investor sentiment. Reports and sentiment data, which are often alphanumeric inputs, are commonly used in this type of research. Additionally, an increasing number of other data sources, such as online traffic figures, consumer transaction data, and satellite imagery, are being used to offer unique insights into market patterns. By integrating and analyzing these diverse datasets, machine learning algorithms can provide more reliable and accurate forecasts regarding financial market movements.

TABLE I. TECHNICAL CALCULATION INDICATORS [17]

Technical Indicators	Calculation Method
Simple Moving Average (MA)	$\frac{C_t + C_{t-1} + \cdots + C_t\text{-}30}{n}$
Exponential Moving Average (EMA)	$EMA(k)_t = EMA(k)_{t-1} + a*(C_t\ EMA(k)_{t-1}$
Momentum (Mom)	$C_t - C_{t-n}$
Stochastic K% (K%)	$\frac{(C_t - LL_{t-n}) \times 100}{HH_{t-n} - LL_{t-n}}$
Stochastic D% (D%)	$\frac{\sum_{I=0}^{n-1} K_{t-I}\%}{n}$
Relative Strength Index (RSI)	$100 - \frac{100}{1 + (\sum_{i+0}^{n-1} Upt - i/n)/(\sum_{I=0}^{n-1} DWt - i/n)}$
Moving Average Convergence/Divergence (MACD)	$MACD(n)_{t-l} + \frac{2}{n+1} * DIFF_t - MACD(n)_{t-1}$
Larry William's R% (LW)	$\frac{H_n - C_t}{H_n - L_n} \times 100$
Commodity Channel Index (CCI)	$\frac{M_t - SM_t}{0,015D_t}$

The technical indicators listed in the table are used as input variables to evaluate the effectiveness of stock market prediction algorithms. Each indicator is computed using a precise mathematical procedure. The Simple Moving Average (SMA) is calculated by averaging the closing prices over a predetermined period. The Exponential Moving Average (EMA) assigns a higher weight to recent prices by smoothing the difference between the previous EMA and the current closing price. Momentum (Mom) measures the rate of change in closing prices over a specific period by subtracting the closing price of a previous period from the current closing price. The Stochastic K% (K%) identifies overbought or oversold conditions by indicating the position of the current closing price relative to the high and low prices over a specified time period. A smoother variation, the Stochastic D% (D%), aids in recognizing potential market reversals. The Relative Strength Index (RSI) evaluates the velocity and direction of market movements by comparing the magnitude of recent gains and losses. The Moving Average Convergence/Divergence (MACD) is a trend-following momentum indicator that illustrates the relationship between two moving averages of an asset's price. Larry William's %R (LW) assesses overbought or oversold conditions by comparing the closing price to the high-low range over a certain time frame. These technical indicators employ distinct mathematical algorithms, rendering them valuable tools for analyzing and predicting stock market trends.

By integrating these indicators, analysts and algorithms can gain a more comprehensive understanding of market behavior and make more informed trading decisions. Additionally, the Commodity Channel Index (CCI) identifies cyclical trends in a commodity by comparing the current price to its moving average and calculating the deviation.

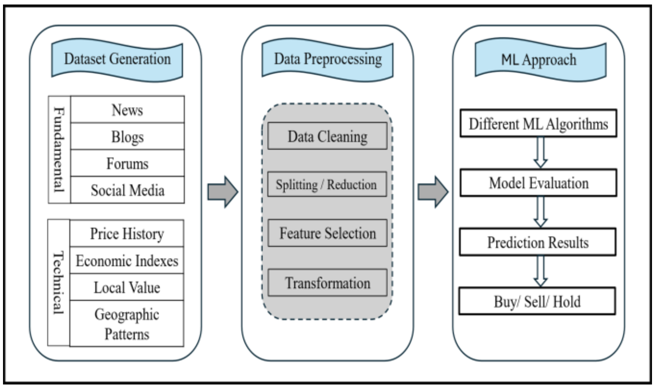

Fig. 1. The framework to predict the movements of the financial market.

Obtaining data for fundamental research is challenging because fundamental indicators are often unstructured. However, with the advancements in AI, it is now possible to use internet data for this purpose, enhancing the accuracy of stock market predictions. This data may include information from an organization's financial reports or insights into investor sentiment. Additionally, the pre-processing stage varies depending on the types of data. In technical analysis, it is crucial for the dataset to normalize the numerical values before using it for the model as a training dataset. Also, data normalization is essential when the ML model aims to identify logical patterns in the provided data. A non-uniform scale of data can lead to inaccurate prediction performance. Therefore, various functions such as RobustScaler, StandardScaler, and MinMaxScaler are used to normalize the data.

IV. Machine Learning Algorithms

Numerous machine learning algorithms have been employed in research projects to forecast stock markets. These can be broadly categorized into two main types: regression models and classification models. Regression models aim to forecast the changes in stock prices, which include a stock's closing price. Classification models, on the other hand, assist investors in making decisions about purchasing, selling, or holding stock. The most frequently employed ML approaches to stock market forecasting are decision trees (DT), SVM, and ANN. In addition to these models, classification strategies in stock market prediction often utilize KNN, RF, Gaussian Naive Bayes (GNB), Bernoulli Naive Bayes (BNB), logistic regression (LR), and XGBoost (XGB). Below is a brief explanation of these algorithms.

A. ARIMA Model

In time series forecasting, the ARIMA model is a widely used statistical technique, particularly for stock price prediction. ARIMA comprises three components: autoregression (AR), which predicts future values using past values; differencing (I), which stabilizes the time series by removing trends and seasonality; and moving average (MA), which uses past forecast errors to make predictions. The parameters of the model are p (the number of lag observations), d (the degree of differencing), and q (the size of the moving average window). The model is denoted as ARIMA(p, d, q). By identifying patterns and trends in historical stock price data, ARIMA forecasts future prices.

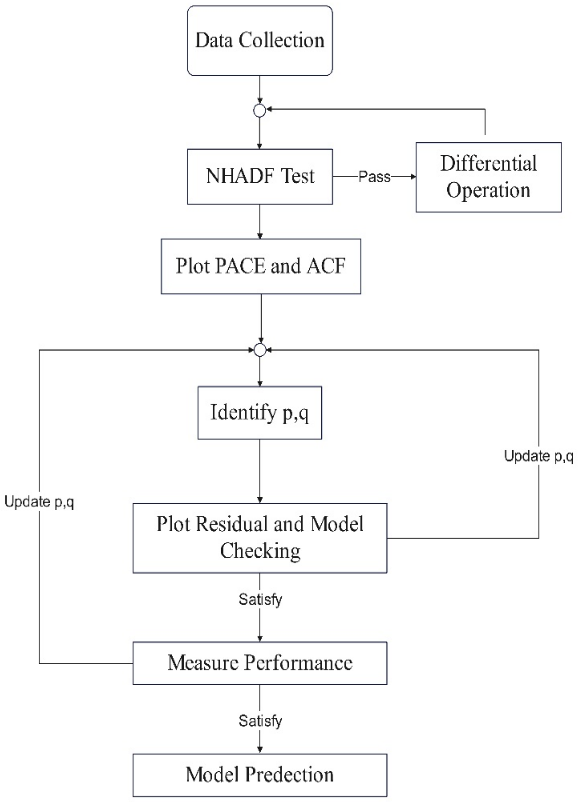

Fig. 2. Working principles of ARIMA model.

Determining the optimal values for p, d, and q involves several key steps: stabilizing the data, fitting the model, and validating it through residual analysis. Autocorrelation and partial autocorrelation plots are used to identify these parameters. While ARIMA is effective, it relies solely on historical data and may require supplementary models or additional data sources to account for unexpected market conditions and external influences.

B. Decision Tree Algorithm

The Decision Tree model is a machine learning technique for stock price prediction that constructs a tree-like structure by splitting the information collected into segments according to the provided value of features Every node in the structure denotes a choice point determined by a certain characteristic, while the branches illustrate the outcomes of these decisions, leading to either further nodes or final predictions at the leaves. When handling continuous data, every interior node in the training data undergoes a logical test $X_i>C$, where a certain threshold level inside the observed spectrum of X_i is represented by C, and X_i is an attribute in the data domain. The threshold volume C is assumed by standards like increasing the descendent nodes' variances or decreasing the overlaps. Considering a dataset with multiple classes represented as C_1, C_2,…, C_T where T signifies the class values, the probability P_i of a class and the sensitivity for every single category is computed. The probability P_i is given by $P_i=\frac{Ci}{T}$. To determine the gain ratio, the following equations needed to be followed sequentially.

$$Entropy\ (T) = -\sum_{i=1}^{n} P_i\ log_2(P_i) \quad (1)$$

$$Gain(B,T) = Entropy\ (T) - \sum_{i=1}^{n} \frac{|T_i|}{|T|}\ (Entropy(T_i) \quad (2)$$

$$Splitting\ Criterion\ (B) = -\sum_{i=1}^{k} \frac{|T_i|}{|T|} Log_2(\frac{|T_i|}{T}) \quad (3)$$

$$Gain\ Ratio = \frac{Gain\ (B,T)}{Splitting\ Criterion\ (B)} \quad (4)$$

C. LSTM

The Long Short-Term Memory is one kind of RNN model that excels in time series forecasting, including stock price prediction. Unlike traditional RNNs, LSTMs are intended to use gating methods to regulate the movement of information in order to identify long-term causalities and trends in data and memory cells that store information. This makes LSTMs particularly effective for analyzing sequential data with temporal dynamics, such as stock prices. By processing input sequences over time, the model learns intricate temporal correlations and patterns. LSTM models utilize large datasets of historical data on stock prices to spot complex trends that conventional statistical methods like ARIMA might overlook.

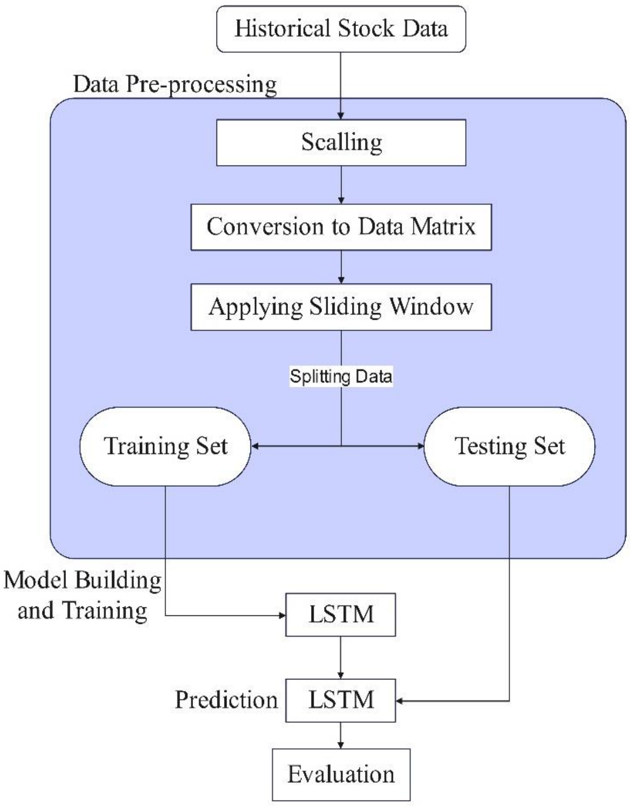

Fig. 3. Working principles of the LSTM model.

The LSTM model modifies the typical ANN approach by replacing the traditional artificial neurons in the buried layer of the network with memory cells. These memory cells effectively manage the association between input and memory, allowing the network to more effectively identify and make use of long-term dependencies in the data.

The LSTM model utilizes both a memory state and a hidden state from the previous time step. The forget gate, controlled by a sigmoid activation function, determines which past information is significant and which can be disregarded, while maintaining relevant information in the memory state. Additionally, the hidden state involves the sigmoid of the current state and the hyperbolic tangent (Tanh) of the memory state, aiding in understanding the significance of the data. The output of the forget gate ranges between 0 and 1 due to the sigmoid activation: an output of 1 indicates full retention of the previous state, 0 signifies complete forgetting, and values in between indicate partial retention of the memory state's relevance.

D. Naive Bayes Algorithm

One of the probabilistic multimodal categorization techniques utilized for predicting the upward and downward movements of stock market indices is the Naive Bayes Algorithm. This method assigns classes to test data by first determining the class to which values from the training set are appropriate. The naive Bayes algorithm forecasts the path of an indicator of the financial economy by leveraging statistical probabilities primarily, notably through Bayes' theorem. Equation 5 represents Bayes' theorem as P(C/X), indicating the probability of event C occurring given that event X has happened.

$$P\left(\frac{C}{X}\right) = \frac{P(C)*(\frac{C}{X})}{P(X)} \quad (5)$$

E. Equations

Artificial Neural Networks (ANNs), which mimic the way human brains process information, is a powerful machine learning technique for stock price prediction. An ANN consists of a network of linked layers made up of a source layer, several layers that are concealed, and an output level of neurons. Each node uses weighted connections and activation functions to detect intricate linkages and trends in the supplied data. ANNs can analyze large volumes of historical data and other relevant financial indicators to capture intricate non-linear relationships for stock price prediction. The network improves its prediction accuracy through an iterative process called backpropagation, where it adjusts the weights of the connections. A fitness function for the ANN algorithm is defined as the square root of the variance in impulses added together for every neuron in the resulting layer. The error signal for the i-th neuron is derived as $e_i=d_i-y_i(k)$ during the k-th training iteration, where y_i is the output value of the i-th neuron, and d_i is the desired value intended for that neuron. The fitness function, represented by Equation (6), is the total sum of the squares of the error signals for all neurons in the ANN's output layer:

$$Fitness\ Function = \sum_i e_i^2 = \sum_i (d_i - y_i(K))^2 \quad (6)$$

This function helps in evaluating the performance of the network by quantifying the discrepancy between real and anticipated values. The fitness function in the ANN algorithm is minimized using the backpropagation technique for error correction. The most precise method to determine the weight values of ANNs, which is essential for updating the fitness function, is the gradient descent approach. The equation for the gradient descent method, where η represents the learning rate, is given by:

$$w_{ij}(k+1) = w_{ij}(k) - \eta \frac{\partial E}{\partial w_{ij}} \quad (7)$$

Here, $w_{ij}(k)$ represents the weight between the i-th and j-th neurons at the k-th iteration, and E is the fitness function, representing the total error. The term $\frac{\partial E}{\partial w_{ij}}$ is the gradient of the error with respect to the weight. By iteratively adjusting the weights using this equation, the network reduces the overall error, improving its prediction accuracy. The backpropagation method is divided into two primary stages: the forward pass and the backward pass. During the forward pass, the model calculates outputs and compares them to the target values. The error rate is then determined by measuring the difference between the actual and desired outputs. In the backward pass, the errors computed during the forward pass are used to adjust the network's weights. These forward and backward operations are repeated multiple times unless the error rate is sufficiently minimized.

V. Model Evaluation Metrics

To evaluate the effectiveness of the prediction strategy during the forecasting process, various evaluation metrics are crucial. These metrics help measure the performance of machine learning algorithms, ensuring their reliability and effectiveness. For classification models, metrics such as the confusion matrix, receiver operating characteristic (ROC) curve, and R-squared are commonly used. These metrics assess the model's ability to correctly classify instances and balance the trade-off between true positive rates and false positive rates.

For regression models, which predict continuous outcomes, metrics like explained variance, mean absolute percentage error (MAPE), root mean square error (RMSE), and mean absolute error (MAE) are employed. These metrics quantify the discrepancies between predicted and actual values, providing insights into the model's precision and accuracy. Explained variance indicates the proportion of the data's variability captured by the model, MAPE measures accuracy as a percentage, RMSE calculates the square root of the average squared differences, and MAE provides the average absolute differences. Understanding and applying these metrics is essential for validating and refining prediction models, ensuring they deliver accurate and reliable forecasts.

A confusion matrix uses a known set of target data to evaluate an ML model's accuracy. Based on this matrix, several other metrics can be derived, including sensitivity, specificity, accuracy, and F1-score. Sensitivity, or recall, measures the probability of correctly identifying a true positive, while specificity indicates the true negative rate. Precision reflects the correctness of the predicted positive classes. The F1-score represents the harmonic mean of precision and recall, balancing the two. Finally, the algorithm's overall accuracy is assessed by evaluating the correctly predicted classes against the total predictions.

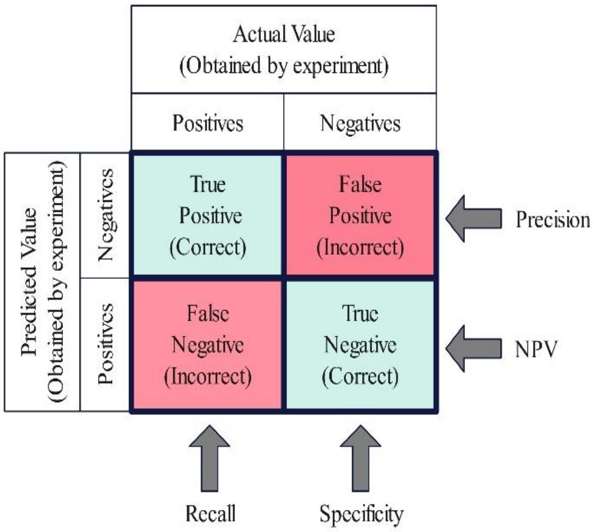

Fig. 4. Confusion matrix concept for the evaluation.

VI. Findings & Discussion

The dataset used in Research [18] to develop a forecasting model grounded upon technical examination was sourced from the "Yahoo Finance" website. It contains historical data for AAPL, representing Apple's stock information over a ten-year period from 2010 to 2021. The dataset comprises 60 features, including moving averages, MACD, RSI, and open, high, and low prices, with the closing price (AAPL's end-of-day price) as the target variable. After identifying the features with the highest correlation to the target, redundant features exhibiting strong connections are combined. Finally, the MinMaxScaler function is applied to scale the data.

To create the ML model, the dataset is divided into three parts: training, validation, and testing data. The model's performance is evaluated using the validation data to fine-tune and adjust its parameters. Finally, the model predicts the target values for the testing dataset, which are then compared with the actual target values. This comparison allows for the assessment of the model's accuracy, using evaluation metrics based on the actual and predicted closing prices.

TABLE II. Performance Analysis of ML Algorithms on 10-Year Apple Stock Price Dataset [7]

Metrics	GNB	XGB	BNB	ANN	DT	RF	LR	KNN	SVM
Precision	0.636	0.71	0.644	0.684	0.62	0.727	0.729	0.684	0.757
Recall	0.634	0.709	0.644	0.684	0.62	0.727	0.727	0.684	0.755
F1-score	0.632	0.709	0.644	0.684	0.62	0.727	0.726	0.684	0.755
Accuracy	0.634	0.709	0.644	0.684	0.62	0.727	0.727	0.684	0.755
AUC	0.63	0.71	0.64	0.68	0.62	0.73	0.73	0.68	0.76

The evaluation metrics and stock price predictions using various algorithms applied to the dataset are summarized in Table II above. According to Table II, the LR model performs better than the LSTM model in predicting AAPL's closing price. Predicting public sentiment using ML algorithms does not yield promising results. The SVM algorithm emerges as the most accurate, with a 76% accuracy rate. Furthermore, the ROC curves illustrate the performance of these methods, showing the AUC for each algorithm, with the SVM algorithm achieving the highest

AUC score. However, several limitations should be considered. Firstly, the comparison between LR and LSTM models is narrow, potentially not reflecting broader model performance across different datasets or time periods. The statement about sentiment analysis yielding unpromising results lacks detail on the specific dataset characteristics or preprocessing steps, limiting generalizability. While SVM achieves a notable 76% accuracy rate and the highest AUC score according to ROC curves, the context of this evaluation—such as validation methods and feature selection—remains unspecified, which could affect the reliability of these metrics. Moreover, the focus on specific algorithms (LR, LSTM, SVM) may overlook potentially superior models, and the temporal stability of predictive models in dynamic markets is not addressed. Further clarification on dataset specifics, validation procedures, and the interpretability of results would enhance the robustness and applicability of these findings in practical settings.

VII. Conclusion

During this study, several ML approaches and their effectiveness were discussed based on Apple's stock information over a ten-year period from 2010 to 2021. Machine learning algorithms can predict stock market changes by analyzing relevant factors like sentiment analysis, trading volumes, historical prices, and economic indicators. It is crucial to choose appropriate models such as SVMs, LSTM networks, decision trees, or linear regression based on the data's characteristics and to adjust their parameters correctly. One of the primary challenges is the natural instability and uncertainty of financial markets, where unforeseen events such as natural disasters, political turmoil, or economic downturns can lead to significant price fluctuations that are difficult to predict. Moreover, the accuracy of these models heavily relies on the comprehensiveness and reliability of the data used for their training. Inaccurate predictions can stem from data that could be more noisy, complete, or biased. Additionally, models that perform effectively on historical data but struggle with new, unfamiliar data often suffer from overfitting. This occurs when models become overly complex, capturing random noise rather than the underlying patterns. However, the effectiveness of employing machine learning algorithms for stock market prediction hinges on various factors and challenges that must be effectively managed. While machine learning (ML) has its limitations, it can still uncover patterns and generate forecasts using extensive datasets and complex models. Moreover, due to the dynamic nature of financial markets, the assumption that historical patterns will reliably forecast future market movements is only sometimes accurate.

Despite these challenges, machine learning (ML) can prove valuable when combined with rigorous data preprocessing, meticulous feature selection, appropriate model choices, robust evaluation techniques, and integration of domain expertise. Successful applications typically adopt a balanced approach, acknowledging both the advantages and limitations of machine learning in stock market forecasting.

References

[1] F. Z. Xing, E. Cambria and R. E. Welsch, "Intelligent Asset Allocation via Market Sentiment Views," in IEEE Computational Intelligence Magazine, vol. 13, no. 4, pp. 25-34, Nov. 2018, doi: 10.1109/MCI.2018.2866727.

[2] Jain, Atishay, Yash Sharma, and Kaushal Kishor. "Prediction and analysis of financial trends using Ml algorithm." Proceedings of the international conference on innovative computing & communication (ICICC). 2021.

[3] Jing, Nan, Zhao Wu, and Hefei Wang. "A hybrid model integrating deep learning with investor sentiment analysis for stock price prediction." Expert Systems with Applications 178 (2021): 115019.

[4] Phuong, Nguyen Thi Thanh, Dinh Tran Ngoc Huy, and Pham Van Tuan. "The evaluation of impacts of a seven factor model on nvb stock price in commercial banking industry in vietnam-and roles of Discolosure of Accounting Policy In Risk Management." International Journal of Entrepreneurship 24 (2020): 1-13.

[5] Ayala, Jordan, Miguel García-Torres, José Luis Vázquez Noguera, Francisco Gómez-Vela, and Federico Divina. "Technical analysis strategy optimization using a machine learning approach in stock market indices." Knowledge-Based Systems 225 (2021): 107119.

[6] Mehtab, Sidra, Jaydip Sen, and Subhasis Dasgupta. "Robust analysis of stock price time series using CNN and LSTM-based deep learning models." 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA). IEEE, 2020.

[7] Hu, Zexin, Yiqi Zhao, and Matloob Khushi. "A survey of forex and stock price prediction using deep learning." Applied System Innovation 4.1 (2021): 9.

[8] Di, Xinjie. "Stock trend prediction with technical indicators using SVM." Independent Work Report, Standford: Leland Stanford Junior University, USA (2014).

[9] A. W. Li and G. S. Bastos, "Stock Market Forecasting Using Deep Learning and Technical Analysis: A Systematic Review," in IEEE Access, vol. 8, pp. 185232-185242, 2020, doi: 10.1109/ACCESS.2020.3030226.

[10] Lu, Wenjie, et al. "A CNN‐LSTM‐based model to forecast stock prices." Complexity 2020.1 (2020): 6622927.

[11] Januschowski, Tim, Jan Gasthaus, Yuyang Wang, David Salinas, Valentin Flunkert, Michael Bohlke-Schneider, and Laurent Callot. "Criteria for classifying forecasting methods." International Journal of Forecasting 36, no. 1 (2020): 167-177.

[12] Rouf, Nusrat, Majid Bashir Malik, Tasleem Arif, Sparsh Sharma, Saurabh Singh, Satyabrata Aich, and Hee-Cheol Kim. "Stock market prediction using machine learning techniques: a decade survey on methodologies, recent developments, and future directions." Electronics 10, no. 21 (2021): 2717.

[13] A. Adebiyi, C. Ayo, A. Marion, O. Sunday O, Stock Price Prediction Using Neural Network with Hybridized Market Indicators, 2012, p. 3. http://www.cisjournal.org.

[14] S. Aktaş, Predıctıon of the Dow Jones Industrial Average and the Nasdaq 100 Indices Usıng Machine Learning, Master Thesis, Social Sciences, 2019. https://polen.itu.edu.tr:8443/server/api/core/bitstreams/2918a720-5dad-4db9-8eeb-ad582bbcb19d/content.

[15] R. Dimingo, J.W. Muteba Mwamba, L. Bonga-Bonga, Prediction of Stock Market Direction: Application of Machine Learning Models, vol. 74, Economia Internazionale/International Economics, 2021. http://www.iei1946.it/upload/rivista_articoli/allegati/425_bonga-et-al-predictionricfinal.pdf. (Accessed 23 January 2023).

[16] A. Subasi, F. Amir, K. Bagedo, A. Shams, A. Sarirete, Stock market prediction using machine learning, Procedia Comput. Sci. 194 (2021) 173–179, https://doi.org/10.1016/j.procs.2021.10.071.

[17] M. Kumar, M. Thenmozhi, Forecasting Stock Index Movement: A Comparison of Support Vector Machines and Random Forest, SSRN Electronic Journal, 2006, https://doi.org/10.2139/ssrn.876544.

[18] Mokhtari, Sohrab, Kang K. Yen, and Jin Liu. "Effectiveness of artificial intelligence in stock market prediction based on machine learning." arXiv preprint arXiv:2107.01031 (2021).