Journal of Mathematics and Modeling in Finance (JMMF)
Vol. 3, No. 2, Summer & Autumn 2023
Research paper

Comparative analysis on forecasting methods and how to choose a suitable one: case study in financial time series

Mahdi Goldani[1]

[1] Faculty of Literature and Humanities, Hakim Sabzevari University, Sabzevar, Iran
mahgoldani@gmail.com

Abstract:
Forecasting in the financial markets is vital for informed decision-making, risk management, efficient capital allocation, asset valuation, and economic stability. This study thoroughly examines forecasting techniques to predict the 30-day closing prices of APPLE in a select group of 100 prominent companies chosen based on their revenue profiles. list of 100 big Companies published by The Fortune Global 500. The evaluated forecasting methods encompass a broad spectrum of approaches, including Moving Average (MA), Exponential Smoothing, Autoregressive Integrated Moving Average (ARIMA), Simple Linear Regression, Multiple Regression, Decision Trees, Random Forests, Neural Networks, and Support Vector Regression (SVR). The information on the dataset was downloaded from Yahoo Finance, and all methods were evaluated in Python. The MAPE method is used to measure the accuracy of the examined methods. Based on the selected dataset, Our findings reveal that SVR, Simple Linear Regression, Neural Networks, and ARIMA consistently outperform other methods in accurately predicting the 30-day APPLE closing prices. In contrast, the Moving Average method exhibits subpar performance, primarily due to its inherent limitations in accommodating the intricate dynamics of financial data, such as trends, seasonality, and unexpected shocks. In conclusion, this comprehensive analysis enhances our understanding of forecasting techniques and paves the way for more informed and precise decision-making in the ever-evolving realm of financial markets. *Keywords:* forecasting methods,

financial assets, time series, Mean Absolute Percentage Error (MAPE).
Classification: MSC2010 or JEL Classifications: C53, G17, C22 .

1 Introduction

The stock market is a cornerstone of modern economies, serving as a crucial avenue for companies to raise capital, enabling individuals to grow wealth, and offering valuable insights into economic health. The stock market is a key driving force

[1]Corresponding author
Received: 02/10/2023 Accepted: 24/12/2023
https://doi.org/10.22054/JMMF.2023.76041.1105

among the various financial markets. A countrys economic situation directly or indirectly impacts sectors such as finance, agriculture, metal, and investment banking. The growth of these sectors hinges on their volatility, which follows the fundamental principle of supply and demand. The demand for a particular sector directly influences the stock market, with increased supply prompting traders and financial institutions to invest in that sector or stock, driving up prices [41]. The ability to predict stock prices is of paramount importance, not only for investors seeking informed decisions but also for policymakers and financial analysts monitoring economic trends. Accurate forecasting in the stock market facilitates efficient allocation of resources, risk management, and informed investment strategies, making it an indispensable tool in the financial world.

Due to the unpredictable nature of the stock market, it is highly difficult for individuals to obtain returns on their investments. Primary, fundamental, and technical analysis are popular approaches to understanding market trends [25], but they possess inherent limitations due to the involvement of lagging indicators and prediction inaccuracy. Forecasting stock market movements is a formidable task due to a myriad of challenges. These include the inherent randomness and unpredictability of market dynamics, non-stationarity of stock data, high volatility, and the influence of investor sentiment and emotions. Additionally, factors like market manipulation, regulatory changes, and rare, unforeseen events add complexity to the process. Balancing model complexity with simplicity and addressing behavioral biases further complicate the task of predicting stock prices.

Within the realm of stock market forecasting, methodologies are distinctly categorized into two overarching branches: quantitative and qualitative methods. In this study, we exclusively focus on the quantitative domain, which harnesses data-driven techniques and statistical models to predict stock prices. These quantitative methods encompass a spectrum of approaches, including time series analysis, machine learning algorithms, neural networks, Monte Carlo simulations, mean-reversion models, volatility models, and factor models. By emphasizing quantitative methods, our research aims to harness the precision and data-driven insights offered by these techniques to predict the closing prices of 100 prominent companies over a 30-day horizon. This deliberate choice underscores our commitment to rigorous numerical analysis and empirical validation in the pursuit of enhancing stock market forecasting accuracy.

This research aims to conduct a comprehensive examination of different forecasting methods applied to the prediction of closing prices for 100 prominent companies, selected based on their revenue. Over a 30-day horizon, we will utilize each forecasting method to estimate the future stock prices of these companies and subsequently assess the accuracy of these predictions. By conducting this empirical study, we intend to shed light on the relative effectiveness of various forecasting techniques in the context of forecasting stock prices for significant corporations. The findings will provide valuable insights for investors, financial analysts, and decision-makers, empowering them to make more informed investment choices and develop robust risk management strategies in the dynamic landscape of the stock market.

2 Literature review

The field of forecasting methods is a widely researched and constantly evolving area in academia. A comprehensive review of the literature reveals a substantial quantity of publications that span several decades. Numerous authors across various fields have contributed to the existing body of literature, resulting in a breadth of knowledge and approach to forecasting methods. In the issue of forecasting, numerous methods and techniques have been developed and explored. The literature encompasses a range of topics, including statistical models, time series analysis, artificial intelligence algorithms, and econometric methods, to name just a few. This diversity in methodologies reflects the interdisciplinary nature of the field and the recognition that no single approach can adequately address the complexity of forecasting in all contexts. The significant multitude of forecasting methods are documented within the realm of academic and professional literature. These diverse approaches have been systematically categorized into two overarching groups. Figure 1 visually presents the spectrum of these methods, offering a comprehensive snapshot of their utilization within the literature.

In summary, the literature on forecasting methods is broad and deep, incorporating a plethora of publications across different disciplines. The field has witnessed a steady growth in the quantity of published research, demonstrating an increasing interest in the advancement of forecasting methodologies. The vibrant citation activity within the field further highlights the influence and impact of key contributions. Overall, this extensive body of literature provides a solid foundation for conducting an in-depth analysis and synthesis of forecasting methods for your forthcoming academic research paper. Some of the key themes that emerge from the literature are as follows:

2.1 Quantitative forecasting methods

This theme focuses on the application of mathematical and statistical models to predict future outcomes. Various techniques such as time series analysis, regression analysis, and machine learning algorithms are commonly employed. The literature discusses the strengths and limitations of these methods, their accuracy in different domains, and the importance of selecting appropriate models and parameters. Some studies debate the trade-off between the complexity and interpretability of quantitative models.

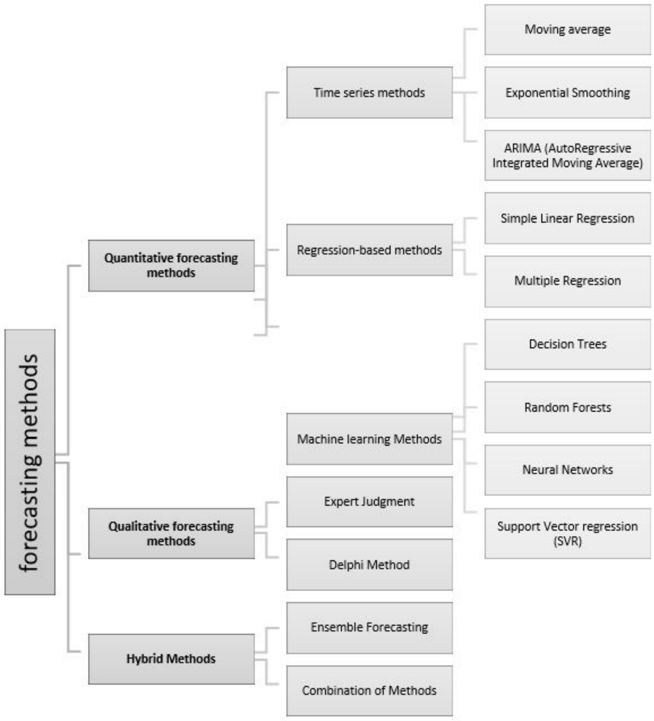

Figure 1: forecasting methods

Time Series Forecasting Methods

The theory of forecasting is based on the premise that current and past knowledge can be used to make predictions. In particular, for time series, there is the belief that it is possible to identify patterns in historical values and successfully implement them in the process of predicting future values [33]. The literature on time series forecasting methods is vast and extensive, demonstrating a significant amount of research activity in this field. Numerous studies have been conducted to develop and evaluate various techniques and approaches for forecasting future values based on historical time series data. Researchers have explored a wide range of methods, including statistical models, machine learning algorithms, and artificial intelligence techniques, to improve the accuracy and reliability of time series forecasting. This collective body of literature reflects the growing interest and importance of time series forecasting in diverse domains such as economics, finance, supply chain management, and environmental sciences. One way to assess the breadth and depth of the available literature is through the number of publications and citation activity. A cursory search reveals a substantial number of research papers, conference proceedings, and books dedicated to time series forecasting methods. Moreover, the citation activity associated with these publications indicates that they have garnered considerable attention from researchers and practitioners alike. The frequent citations suggest that the literature is not only expanding in terms of volume but also attracting interest and being integrated into the broader body of knowledge. This theme focuses on the use of statistical models and techniques for time series forecasting. Various statistical methods, such as ARIMA, exponential smoothing, and Moving average, are explored and compared. The debate within this theme revolves around the relative performance and limitations of these statistical methods in different forecasting scenarios. Hyndman [21] introduce a state space framework for automatic forecasting using exponential smoothing. The state space framework for automatic forecasting using exponential smoothing methods provides a unified approach that is flexible and effective. It allows for accurate parameter estimation and reliable forecast computation, making it a valuable tool for time series forecasting. Taylor [42] focuses on the evaluation of methods for very short-term time series forecasting, typically involving forecasts up to a few hours ahead. In this article, the performance of various methods is compared, including exponential smoothing, ARIMA models, and artificial neural networks, on multiple real-world datasets. The paper discusses the evaluation metrics used, the forecasting accuracy of different methods, and the computational requirements. It also explores the impact of data frequency and the availability of exogenous variables on the forecasting performance. The main conclusion of this paper is that the choice of forecasting method for very short-term time series forecasting depends on the specific characteristics of the data and the desired level of accuracy. The author finds that simple exponential smoothing methods often perform well in these scenarios, particularly when the data exhibit regular patterns. However, more complex methods such as ARIMA models and artificial neural networks can provide better accuracy when the data exhibit non-linear or complex patterns. The paper highlights the importance of considering computational requirements and the availability of exogenous variables when selecting a forecasting method. Hansun [18] has introduced a new approach to the moving average method in time series analysis. The proposed approach uses the basic formula of Weighted Moving Average (WMA) to get a base value, which will then be used to get the forecasted value using Exponential Moving Average (EMA) formula. experimental result showed that the weighted exponential moving average method can be used to forecast the JKSE composite index data, which is an indicator of stock price changes in Indonesia. In a study by [29], the authors modeled and forecasted daily air temperature and precipitation time series from four European sites with different climatic zones. They employed various forecasting methods, including Box-Jenkins, Holt-Winters seasonal auto regressive integrated moving-average, autoregressive integrated moving-average with external regressors, and time series regression. The results demonstrated that the models effectively captured the dynamics of the time series data and generated reliable forecasts. In a study by [19], a comparison of forecasting methods was conducted using new student admission data in a university study program. The single, double, and triple exponential smoothing methods were evaluated. The best forecasting accuracy was achieved with the triple exponential smoothing method, yielding a mean percentage error (MPE) of 0.0161, using parameter values of $\alpha = 0.6$ and $\beta = 0.9$. This research underscores the effectiveness of the triple exponential smoothing method in predicting future student admissions. In a study by [46], the authors proposed a hybrid model for stock price forecasting, called DWT-ARIMA-GSXGB. The model utilizes the discrete wavelet transform to split the data set into approximation and error parts, which are then processed by ARIMA models and an improved xgboost model (GSXGB), respectively. The prediction results are combined using wavelet reconstruction. The experimental comparison of 10 stock data sets showed that the DWT-ARIMA-GSXGB model outperformed the four prediction models of ARIMA, XGBoost, GSXGB, and DWT-ARI-MA-XGBoost in terms of error. The simulation results demonstrated that the proposed model has good approximation and generalization abilities, and can fit the stock index opening price well. This research highlights the potential of the DWT-ARIMA-GSXGB model to improve the predictive performance of a single ARIMA or XGBoost model in stock price forecasting. [11] developed a web-based sales forecasting system for UD Parama Store,. The system utilized the double exponential smoothing brown method, which involves smoothing and exponential decrease to handle linear and non-stationary data. The forecasting accuracy was evaluated using the MAPE method, resulting in error rates ranging from 7.99% to 32.42% for 10 different items.

Regression Models

In the study conducted by [46] , the MLR algorithm was employed for attribute weight calculation, effectively addressing attribute redundancy. A novel approach was introduced, presenting a weighted naive Bayesian algorithm derived from multiple regression, termed MLWNBC (Multiple Linear Weighted Naive Bayesian Classifier). This approach assigns weights to attributes, quantifying their impact. By incorporating weights, MLWNBC refines the WNBC (Weighted Naive Bayesian Classifier), making the classification process more rational. Experimental results from classifying 10 datasets within the UCI database demonstrated the algorithm's robust capabilities, showcasing improved accuracy and reduced processing time. It's worth noting that some attributes, as determined by data collection, exhibit negligible influence on the outcomes. [32] It is proposed to use a multivariate statistical method, i.e. factor analysis, to identify predictor variables by their relationships and importance, to approximate portfolio sensitivities to 4 chosen macroeconomic factors (Market Performance, Real GDP, Inflation, and Unemployment). (Market Performance, Real GDP, Inflation, and Unemployment). Introduces and applies a multi-factor model for portfolio management of stocks. First, the model is established, the portfolio will then be refined and multi-factors will eventually be used to estimate portfolio sensitivity. Results show that improved results can be obtained by choosing the less associated variables. In the work by [17], a framework was created to employ linear regression for assessing software characteristics within defined applications. This approach utilized code vectors derived from software instructions. Through conducted experiments, it was indicated that linear regression could serve as an effective technique for categorizing interconnected software during software analysis. In summary, a thoughtfully designed machine-learning model can seamlessly contribute to software analysis, further bolstering the understanding of software functionality through the incorporation of machine learning into information analysis.

Machin learning

Voyant [44] examine regression trees, random forests, and gradient-boosting machine-learning methods for solar radiation forecasting. As a conclusion of this paper, the ANN and ARIMA methods are equivalent in terms of the quality of prediction. Makridakis [26] in a study titled Statistical and Machine Learning Forecasting Methods: Concerns and Ways Forward, show ML methods need to become more accurate, require less computer time, and be less of a black box. A major contribution of this paper is in showing that traditional statistical methods are more accurate than ML ones and pointing out the need to discover the reasons involved, as well as devising ways to reverse the situation. Wu [48] proposed DNN-BTF, a Deep Neural Network model for traffic flow prediction using big data. Unlike existing models, DNN-BTF incorporates spatial-temporal characteristics and weekly/daily periodicity. It employs attention-based learning for past traffic flow significance, integrates a Convolutional Neural Network for spatial features, and utilizes Recurrent Neural Network for temporal features. The model challenges the "black-box" perception of neural networks in transportation by providing visualizations. Validated on PeMS database, DNN-BTF outperformed state-of-the-art methods in long-term traffic flow prediction tasks. Zeroua [49] in their article conducted a comparative study on forecasting COVID-19 cases using daily data from Italy, Spain, France, China, USA, and Australia. Employing five deep learning models (RNN, LSTM, BiLSTM, GRUs, and VAE), the study trained and evaluated the models on the available data. Results indicated that the Variational AutoEncoder (VAE) outperformed other models, demonstrating superior accuracy in predicting COVID-19 cases. The study's global perspective emphasized the adaptability of these deep learning models to diverse COVID-19 dynamics, offering insights for effective short-term forecasting in the ongoing pandemic.

2.2 Qualitative forecasting methods

Qualitative forecasting methods emphasize subjective judgment, expert opinions, and qualitative data to forecast future trends. This theme explores techniques such as the Delphi method, scenario planning, and market research-based forecasting. The literature highlights the value of expert knowledge, the role of human judgment in decision-making, and the challenges of effectively utilizing qualitative data. Debates within this theme revolve around the objectivity and reliability of qualitative approaches compared to quantitative methods.

Central to this discourse is the resounding conclusion that robust forecasting hinges upon objective and scientific underpinnings, steering clear of subjective conjectures. Practical guidelines for refining accuracy underscore the need for transparency and accountability throughout the forecasting journey.

Turning to Chris Chatfield's [9] *"Time-Series Forecasting"*, a distinct spotlight shines on time-series forecasting methods. This paper embarks on a deep dive into various time-series models, including exponential smoothing, ARIMA models, and state space models. Echoing through its pages is the need for astute model selection, validation, and comprehension of the intricacies embedded within time-series data.

The resounding conclusion underscores that time-series forecasting's efficacy rests upon meticulous model selection and validation, an intricate dance that demands an adept understanding of underlying patterns and dynamics. This echoes the call for perpetual vigilance in the form of ongoing model monitoring and adaptation.

In a distinct realm, Fildes and Kourentzes [15] research paper *"Validation and Forecasting Accuracy in Models of Climate Change"* navigates the intricate intersections of forecasting models and climate change. The authors expound on the intricacies and challenges of climate change forecasting, rigorously assessing the performance of diverse models. Amid the discourse, validation techniquesencompassing out-of-sample testingtake center stage, as does the insistence on robust, reliable models within the domain of climate change research.

From this exploration surfaces the resolute assertion that precise climate change forecasting necessitates validated, robust models. The paper sheds light on the gaps within existing models when confronting the complexities of climate change, extolling the need for further research and developmental strides. Transparency and reproducibility stand as steadfast values, underpinning the integrity of climate change forecasting.

Collectively, these works enrich the fabric of forecasting wisdom, illuminating an array of techniques while offering profound insights into the application and refinement of predictive models across diverse domains. From the fusion of qualitative and quantitative approaches to the rigorous underpinning of climate change forecasting, these works collectively pave the way for refined, evidence-based forecasting practices.

2.3 Combining quantitative and qualitative methods

The theme explored in this context revolves around the symbiotic relationship between quantitative and qualitative forecasting methods. Recognizing the potential synergies arising from their integration, both researchers and practitioners have engaged in a discourse to harness the strengths of each approach while mitigating their respective limitations. This dialogue is marked by the exploration of hybrid forecasting frameworks that intertwine judgment-based insights with statistical methodologies, as well as the infusion of qualitative perspectives into quantitative models. Within this conversation, debates ensue regarding the most effective amalgamation strategies, the challenges posed by integration, and the ensuing impact on both forecast accuracy and decision-making processes.

"Principles of Forecasting: A Handbook for researchers and practitioners" by Armstrong [2] constitutes an exhaustive compendium elucidating forecasting principles and practices. The handbook traverses diverse terrain, including forecasting accuracy evaluation, extrapolation methodologies, judgmental forecasting, and the integration of forecasting into decision-making. Drawing insights from empirical evidence, the author dispenses pragmatic recommendations. Foremost among the handbook's conclusions is the imperative of evidence-based forecasting grounded in scientific principles and meticulous evaluation. Furthermore, the work accentuates the significance of transparency, accountability, and validation within the forecasting domain, while also underscoring the avoidance of common pitfalls like overfitting models and overlooking base rates.

"Time-series forecasting" by Chatfield [9] places a spotlight on time series forecasting, a widely employed approach for prognosticating future values based on historical data. The author delves into various time series models, encompassing exponential smoothing, ARIMA models, and state space models. This exploration extends to practical considerations such as model selection, parameter estimation, and model diagnostics. The principal conclusion encapsulated within this work is the meticulous nature of time series forecasting, demanding comprehensive modeling and analysis of the underlying dynamics and patterns within the data. The work underscores the need to grasp the limitations of distinct time series models, while also emphasizing the continual monitoring and updating of forecasts as new data emerges.

Meneghini et al. [29] present a novel demand forecasting method that integrates quantitative models with qualitative contextual factors. The approach involves selecting the best-fitting mathematical model based on Rš and MAPE from historical data. Following this, forecasts generated by the chosen model are adjusted through expert judgement, incorporating contextual factors like events or renovations not present in the historical data. Applied to a fast-food restaurant's meat demand forecasting, the adjusted method demonstrated an average error of 10% in the worst scenario, significantly outperforming the quantitative model without judgemental adjustment, which resulted in an average error of 38%. This innovative method offers improved accuracy by combining quantitative modeling with expert insights into contextual factors.

In Khemavuk and Leenatham's [23] research, the focus is on introducing new demand forecasting concepts using artificial intelligence methods. The study tracks the evolution of forecasting from traditional to AI methods, specifically employing ANN and ANFIS with both quantitative and qualitative data. Model structures are detailed for single and combined forecasting methods to enhance accuracy. Two key research questions are addressed: the effectiveness of proposed methods with qualitative data compared to those without, and whether a combined forecasting method outperforms a single method. This work contributes to advancing demand forecasting methodologies, providing insights valuable for decision-makers.

Kucharavy et al. [24] address the challenge of reinforcing strategic decision-making through reliable forecasts of technological change. The study observes that existing strategic forecasts heavily rely on expert opinions, which are prone to cognitive biases. To mitigate this, the Researching Future method (RFm) is introduced, offering a mixed methods approach combining problem-based strategies and logistic functions within an applied resources paradigm. The article presents a practical case study demonstrating RFm's application, highlighting its potential to manage cognitive biases in technology forecasting. The study contributes to technology forecasting methodology, particularly benefiting specialists in copper mining technology R and D, among others.

2.4 Forecasting methods in finance database

A literature review on forecasting methods in finance databases reveals a rich and dynamic field that combines traditional quantitative techniques with cutting-edge data analytics and machine learning approaches. Researchers and practitioners in finance have been continuously exploring and refining these methods to improve forecasting accuracy, risk management, and investment strategies.

The shortage of comprehensive literature comparing forecasting methods in the financial market underscores a significant gap in the current body of knowledge. While numerous studies may individually explore different forecasting techniques, there is a distinct lack of holistic examinations that systematically compare and contrast these methods. This gap in the literature raises questions about the effectiveness and applicability of various forecasting approaches in the dynamic and complex world of finance. To address this shortcoming, this research aims to bridge this gap by conducting a thorough comparative study that evaluates the strengths, weaknesses, and practical implications of different forecasting methods within the finance sector.

3 Methodology

Firstly, the list of 100 big Companies published by The Fortune Global 500 is prepared. The Fortune Global 500, also known as the Global 500, is an annual ranking of the top 500 corporations worldwide as measured by revenue. The list is compiled and published annually by Fortune magazine . The close price of each

Table 1: Literature review in finance forecasting methodology

	auther('s)	Main finding
forecasting methods	Bijesh et al. [12]	The ARIMA model's predicted prices closely aligning with actual prices on the specified dates signifies its successful implementation and accuracy in handling the time series data.
	Meher et al [28]	ARIMA model estimation selected top 5 models with varied AR and MA terms, adjusted based on Volatility, R-squared, and AIC for insightful stock price prediction analysis.
	Du [14]	ARIMA-BP neural network surpasses BP neural network in prediction accuracy, while BP neural network outperforms linear ARIMA model, confirming the nonlinearity in stock price index changes.
MA	Naik and Mohan [31]	Moving average statistics are considered to identify the stock crisis points.
	Shah and Isah. [37]	Implemented Echo State Networks (ESN), a subclass of RNN, to predict S and P 500 stock prices using price, moving averages, and volume as features.
Exponential Smoothing	Sidqi and Sumitra [40]	The result showed that the MAPE of Single Exponential Smoothing is 20% and the MAPE of Double Exponential Smoothing is around 24%.
	Shukor et al. [38]	Based on the analysis done, the result shows that Holt's Linear Trend is the better forecasting method compared to Double Exponential Smoothing and Random Walk.
Machine Learning	Chatzis et al. [10]	Utilizing Deep Neural Networks substantially enhances classification accuracy, providing a robust and efficient global systemic early warning tool surpassing established methods.
	Khan et al. [22]	Random forest classifier is found to be consistent and the highest accuracy of 83.22% is achieved by its ensemble.
	Ghanbari and Arian. [16]	BOA-SVR algorithm performs as one of the best models among the total set of twelve methods studied
	Chandwani et al. [8]	Integrating Genetic Algorithms enhances ANN accuracy, and combining technical analysis with SVM and ANN proves effective for Indian stocks, facilitating investors and traders in maximizing quarterly profits.

company is downloaded from the Yahoo finance library in Python. The period of Data is 4 January 2018 to 22 August 2023 as daily, and only the data of companies with complete data are acceptable.30 last day of APPLE close price is removed to evaluate the accuracy of each method in this dataset. The Fig 2 shows the process of this article.

3.1 Time series methods

Moving Average

Forecasting future trends and patterns in data is an essential endeavor across various industries and disciplines. As businesses aim to make informed decisions, economists analyze economic indicators, and researchers study evolving phenomena, the ability to predict future outcomes becomes a crucial asset. In this pursuit, "Moving Averages" emerge as a fundamental tool in time series analysisa field that deals with data points ordered chronologically. Moving averages provide insights into trends, patterns, and underlying structures within noisy and fluctuating data. Time series data often exhibit short-term fluctuations that can obscure the underlying trends. These fluctuations could result from various factors such as seasonality, random noise, or transient events. Analyzing raw data points might lead to misinterpretations, hindering our ability to identify true patterns. This is where moving

Figure 2: forecasting process steps

averages come into play: they allow us to smoothen out the noise and focus on the overarching trends, making the data more manageable and interpretable. The main advantages of the moving average stock level indicator is that it offers a smooth line and also helps to cut down the amount of noise on the price chart compared with other level of indicators [4]. The moving average is calculated by adding the closing price and then dividing this total by the number of periods. The following Eq. (1) is applied for the moving average,

$$f_t = \frac{A_{t-1} + A_{t-2} + A_{t-3} + \ldots + A_{t-n}}{n} \tag{1}$$

f_t = forecast for the coming period
A_{t-1} = Actual occurrence in the past period for up to n periods
N = number of periods to be averaged.

Exponential smoothing

Exponential smoothing was first suggested in the statistical literature without reference to previous work by Robert Goodell Brown in 1956 and then expanded by Charles C. Holt in 1957. Exponential smoothing is a broadly accurate principle for smoothing time series data using the exponential window function. The controlling input of the exponential smoothing calculation is defined as the smoothing factor or the smoothing constant. The simplest form of an exponential smoothing formula is given by:

$$s_t = \alpha x_t + (1-\alpha)s_{t-1} = s_{t-1} + \alpha(x_t - s_{t-1}) \tag{2}$$

s_t = smoothed statistic, it is the simple weighted average of the current observation
x_t = previous smoothed statistic
α = smoothing factor of data; $0 < \alpha < 1$
t = period
If the value of the smoothing factor is larger, then the level of smoothing will reduce. A value of α close to 1 has less of a smoothing effect and gives greater weight to recent changes in the data, while the value of α closer to zero has a greater smoothing effect and is less responsive to recent changes.

Auto-Regressive Integrated Moving Average (ARIMA)

ARIMA models are, in theory, the most general class of models for forecasting a time series which can be made to be stationary by differencing (if necessary), perhaps in conjunction with nonlinear transformations such as logging or deflating (if necessary). A random variable that is a time series is stationary if its statistical properties are all constant over time. A stationary series has no trend, its variations around its mean have a constant amplitude, and it wiggles in a consistent fashion, i.e., its short-term random time patterns always look the same in a statistical sense. The latter condition means that its autocorrelations (correlations with its own prior deviations from the mean) remain constant over time, or equivalently, that its power spectrum remains constant over time. A random variable of this form can be viewed (as usual) as a combination of signal and noise, and the signal (if one is apparent) could be a pattern of fast or slow mean reversion, sinusoidal oscillation, or rapid alternation in sign, and it could also have a seasonal component. An ARIMA model can be viewed as a filter that tries to separate the signal from the noise, and the signal is then extrapolated into the future to obtain forecasts. In terms of y, the general forecasting equation is:

$$\hat{y}_t = \mu + \phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} - \theta_1 e_{t-1} - \ldots - \theta_q e_{t-q} \tag{3}$$

Here the moving average parameters θs are defined so that their signs are negative in the equation, following the convention introduced by Box and Jenkins. Some authors and software (including the R programming language) define them so that they have plus signs instead. When actual numbers are plugged into the equation, there is no ambiguity, but its important to know which convention your software uses when you are reading the output. Often the parameters are denoted there by AR(1), AR(2), , and MA(1), MA(2), etc..

3.2 Regression methods

Regression serves two main purposes. Firstly, it is commonly employed for forecasting and prediction, aligning closely with the domain of machine learning. Secondly, regression analysis can also be harnessed to establish causal connections between independent and dependent variables in certain scenarios. It's important to note that regression, on its own, reveals relationships solely between a dependent variable and a predetermined dataset containing various variables [47]. Regression models indicate that the independent variables are used to anticipate the values of dependent variables. Through regression analysis, the estimated value of the dependent variable 'y' is determined based on a spectrum of values for the independent variable 'x' [34]. Our paper delves into the exploration of linear regression and polynomial regression, which offer improved fitting for predictive modeling. Regression can manifest as either a straightforward linear regression or a more complex multiple regression scenario.

Simple Regression

Simple Linear Regression, as depicted by a model with a solitary independent variable , establishes the relationship of dependence for a variable. The equation 4 embodies this relationship in simple linear regression. This technique enables the isolation of independent variable effects from dependent variable interactions, enhancing the clarity of their respective impacts [1].

$$y = \beta_0 + \beta_1 x + \epsilon \tag{4}$$

Multiple Linear Regression

Multiple Linear Regression (MLR) stands as a statistical method employed to forecast the outcome of a response variable by considering multiple explanatory variables. The core aim of MLR is to create a model capturing the linear association between the independent variables (x) and the dependent variable (y) under examination [43]. The fundamental structure of the MLR model is articulated as follows:

$$y = \beta_0 + \beta_1 x_1 + \ldots + \beta_m x_m + \epsilon \tag{5}$$

The formulation for determining the coefficient matrix is delineated by the equation 6:

$$\hat{\beta} = (X^T X)^{-1} X^T y \tag{6}$$

where:

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_m \end{bmatrix}, X = \begin{bmatrix} 1 & x_1 1 & x_1 2 & \ldots & x_1 m \\ 1 & x_2 1 & x_2 2 & \ldots & x_2 m \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 1 & x_n 1 & x_n 2 & \ldots & x_n m \end{bmatrix}, y = \begin{bmatrix} y_0 \\ y_1 \\ \vdots y_n \end{bmatrix} \tag{7}$$

3.3 Machine learning methods

Neural network models

Artificial neural networks are forecasting methods that are based on simple mathematical models of the brain. They allow complex nonlinear relationships between the response variable and its predictors. A neural network can be thought of as a network of neurons which are organized in layers. The predictors (or inputs) form the bottom layer, and the forecasts (or outputs) form the top layer. There may also be intermediate layers containing hidden neurons Basic neural networks consist of input and output layers without any hidden layers, resembling linear regressions. In Figure 3, the neural network representation of a linear regression

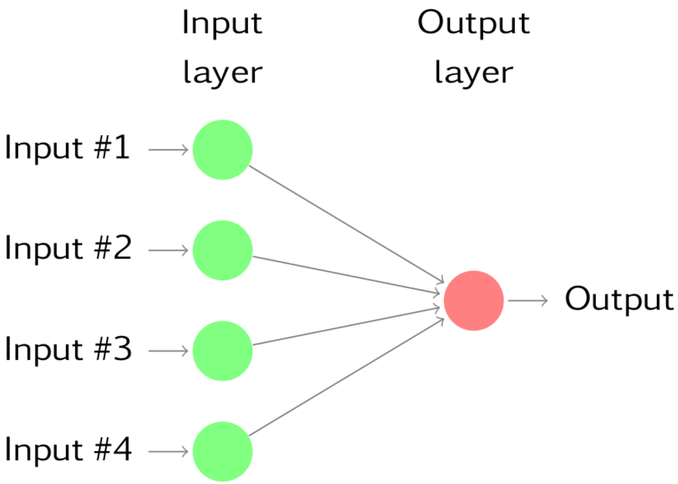

Figure 3: Neural network models

with four predictors is depicted. The coefficients linked to these predictors are referred to as "weights." The predictions are computed through a linear combination of the inputs. These weights are determined in the neural network context through a "learning algorithm" that minimizes a "cost function," often the Mean Squared Error (MSE). However, for this uncomplicated scenario, linear regression can be employed as a more efficient approach for model training. This is known as a multilayer feed-forward network, where each layer of nodes receives inputs from the previous layers. The outputs of the nodes in one layer are inputs to the next layer. The inputs to each node are combined using a weighted linear combination. The result is then modified by a nonlinear function before being output.

$$z_j = b_j + \sum_{i=1}^{4} w_{i,j} x_i \tag{8}$$

In the hidden layer, this is then modified using a nonlinear function such as a sigmoid,

$$s(z) = \frac{1}{1 + e^{-z}} \tag{9}$$

To give the input for the next layer, this tends to reduce the effect of extreme input values, thus making the network somewhat robust to outliers. The parameters b_1, b_2, b_3, and $w_{1,1}, \ldots, w_{4,3}$ are "learned" from the data. The values of the weights are often restricted to prevent them from becoming too large. The parameter that restricts the weights is known as the "decay parameter" and is often set to be equal to 0.1. The weights take random values to begin with, and these are then updated using the observed data. Consequently, there is an element of randomness in the predictions produced by a neural network. Therefore, the network is usually trained several times using different random starting points, and the results are averaged.

Decision tree

The methodology of decision trees involves a recursive process of splitting data based on the most significant attributes (features) to make decisions or predictions. The most commonly used algorithm for constructing decision trees is the CART (Classification and Regression Trees) algorithm, which is based on a measure of impurity such as Gini impurity for classification problems and mean squared error for regression problems. At each step, the algorithm evaluates all possible attributes and their values to determine the best way to split the data. It measures how well a split separates the data into distinct classes or values. The choice of splitting criteria depends on the type of problem:

- For classification (predicting categories or classes), Gini impurity and Entropy are commonly used. The Gini impurity formula is:

$$Gini(D) = 1 - \sum_{i-1}^{c} p_i^2 \tag{10}$$

  where D represents the dataset, c is the number of classes, and p_i is the proportion of samples belonging to class i in the dataset.

- For regression (predicting numerical values), Mean Squared Error (MSE) is frequently used. The MSE formula is:

$$MSE(D) = \frac{1}{|D|} \sum_{i \in D} (y_i - \hat{y}_i)^2 \tag{11}$$

  where D is the dataset, $|D|$ is the number of samples in the dataset, y_i is the target value for sample i, and $\bar{y}$ is the mean target value for all samples in D.

The algorithm continues to split the data into subsets based on the selected attribute and value until a stopping condition is met. This condition could be a maximum depth limit, a minimum number of samples in a leaf node, or other user-defined criteria. Once a stopping condition is met, a leaf node is created, which represents a class (for classification) or a predicted value (for regression).After building the tree, a pruning step may be applied to remove branches that do not contribute significantly to the model's performance. Pruning helps prevent overfitting. To make predictions on new data, you follow the branches of the decision tree from the root node to a leaf node, and the class or value associated with that leaf node is the prediction. This is a simplified overview of the methodology. Decision tree algorithms can be further refined and extended, and there are variations such as Random Forests and Gradient Boosted Trees that enhance their performance and robustness.

Random Forest (RF)

Random Forest (RF) is a collection of regression trees, each specified in a bootstrap sample of the original data. The method was originally proposed by Breiman [7].

Since we are dealing with time series, we use a block bootstrap. Suppose there are B bootstrap samples. For each sample b, a tree with Kb regions is estimated for a randomly selected subset of the original regressors. Kb is determined in order to leave a minimum number of observations in each region. The final forecast is the average of the forecasts of each tree applied to the original data:

$$\hat{y}_{t+\frac{h}{t}} = \frac{1}{B}\sum_{b=1}^{B}\left[\sum_{i=1}^{T_b}\hat{\beta}_{i,b}B_{Ji,b}(X_t;\hat{\theta}_{i,b})\right] \tag{12}$$

The theoretical foundation of Random Forest (RF) models has primarily been established for scenarios involving independent and identically distributed random variables. For instance, Scornet [35] demonstrated the consistency of the RF approximation in estimating the unknown function fh(Xt). In a more recent study, Wager and Athey [45] provided evidence of the RF estimator's consistency and asymptotic normality.

Support Vector Machine (SVM) Algorithm

SVR was developed by Vapnik and co-workers [13] by extending their SVM algorithm for classification [5]. Support Vector Regression (SVR) is a machine learning algorithm used for regression tasks. It aims to find a function that fits the training data while minimizing the margin of error. The formula for SVR can be described as follows: For linear SVR, the basic formula is similar to that of a linear regression model, but with the introduction of a margin of tolerance ε around the regression line to account for deviations:

$$y = f(x) = wx + b = w^T x + B \tag{13}$$

$$\min \frac{1}{2}\|W\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \tag{14}$$

Subject to:

$$\begin{aligned} &y_i - w^T x_i - b \leq \varepsilon \\ &w^T x_i + b - y_i \leq \varepsilon \end{aligned}$$

These formulas represent the optimization problem that SVR aims to solve to find the best-fitting function for regression tasks. The specific kernel function used can vary, with popular choices including linear, polynomial, radial basis function (RBF), and sigmoid kernels, among others. The choice of kernel depends on the nature of the data and the problem at hand [50].

3.4 Forecast performance measures

The accuracy of forecasts indicates how well a forecasting model predicts the chosen variable. one of the most common methods used to calculate Forecasting Accuracy is MAPE which is abbreviated as Mean Absolute Percentage Error. It is an effective and more convenient method because it becomes easier to interpret the accuracy just by seeing the MAPE value. The formula to calculate MAPE is:

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{A_i - F_i}{A_i}\right| \tag{15}$$

4 Result

Before applying the forecasting methods, we divided the 1417 data into the training set and the test set; we used 1387 of the data for training the models and the remaining 30 for testing the models. ARIMA models have three key parameters: the order of autoregression, the degree of differencing, and the order of the moving average. These parameters are represented as p, d, and q, respectively. According to the pmdarima.arima library the best model of ARIMA for this database is ARIMA(0,1,1). In the context of using regression for prediction, it is imperative to assess the stationarity of the variables involved. In this study, the Dickey-Fuller test was employed for this purpose. According to the test results, it was observed that none of the variables were stationary at their original levels; however, they all became stationary after undergoing first-order differentiation. In this article, we aim to forecast the 30-day closing price of Apple Inc. (AAPL) stock. We'll use nine different forecasting methods to analyze historical data and gain insights into where AAPL's stock price might be headed. Table 2 includes the results of these methods.

Each technique possesses certain advantageous characteristics which should be suitably applied to the case. This section throws light on the advantages and disadvantages of each for the techniques discussed previously. The fig 4 shows the predicted values using these techniques and the actual values, which shows how well each of these methods have been able to make a good prediction from this database.

Mean Absolute Percentage Error method is used to measure the accuracy of each. The table below shows the results of the predictive accuracy of each method.

SVR, Simple Regression, Neural network and ARIMA have best performance to predict 30 days of close price in this dataset. SVR, a powerful technique, was great at handling both simple and tricky relationships in the data. It was also good at dealing with unusual data points and could be adjusted to work better with different settings. Simple Linear Regression, despite being basic, was handy due to its simplicity and quick setup, especially when the relationship between variables was

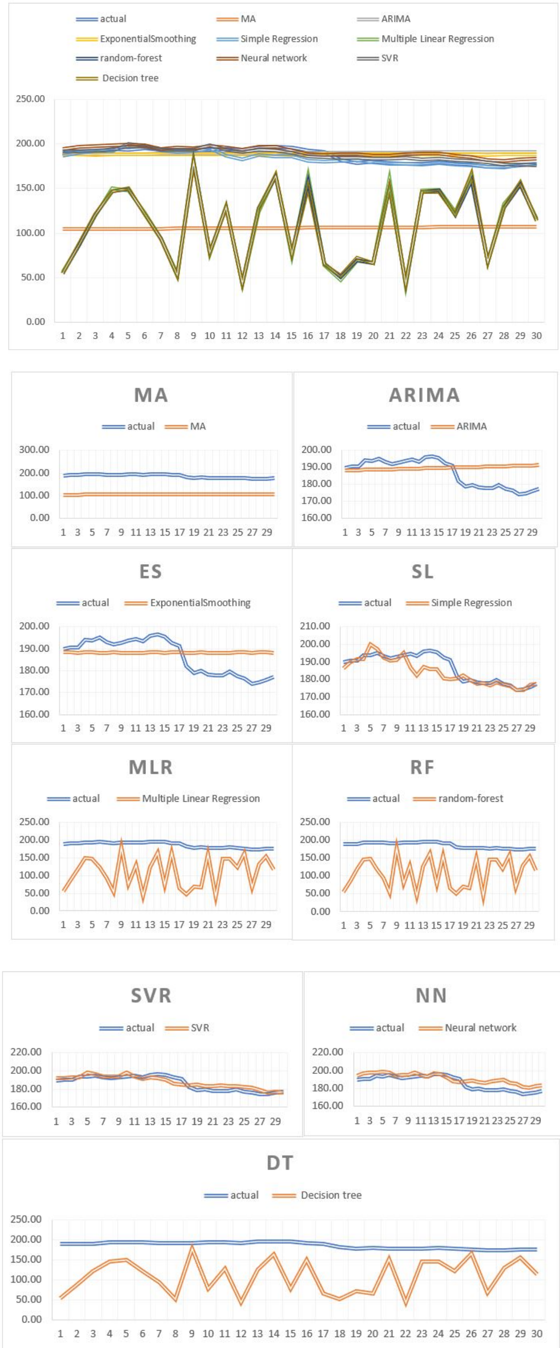

Figure 4: Comparison of the results of various forecasting methods

Table 2: result of forecasting methods

	Actual	MA	ARIMA	ES	Simple Regression	Multiple Linear Regression	Random Forest	Neural Network	SVR	Decision Tree
1	189.77	104.29	188.23	188.42	186.49	56.79	54.93	194.21	191.82	54.69
2	190.54	104.40	188.33	188.28	189.83	85.98	86.45	196.56	191.97	88.21
3	190.69	104.50	188.44	188.05	191.41	119.55	120.28	197.86	192.91	121.26
4	193.99	104.61	188.54	188.40	191.71	149.96	147.01	198.17	193.10	146.63
5	193.73	104.72	188.65	188.29	200.13	147.73	149.23	199.04	197.89	150.04
6	195.10	104.83	188.75	188.22	197.43	122.58	120.87	198.12	196.49	121.21
7	193.13	104.94	188.86	188.08	192.40	93.96	94.08	194.41	193.43	93.25
8	191.94	105.04	188.96	188.40	190.51	54.00	52.28	195.48	193.73	52.25
9	192.75	105.15	189.07	188.25	191.33	180.82	179.76	194.78	193.77	178.19
10	193.62	105.26	189.17	188.17	194.92	76.33	79.00	197.93	198.55	77.83
11	194.50	105.37	189.27	188.17	186.83	130.73	129.04	195.53	193.39	128.98
12	193.22	105.47	189.38	188.11	182.52	42.12	43.72	193.48	191.05	43.59
13	195.83	105.58	189.48	188.42	187.20	123.77	127.45	197.00	193.27	126.66
14	196.45	105.69	189.59	188.28	185.70	164.20	163.42	196.33	192.18	165.12
15	195.61	105.80	189.69	188.05	185.96	74.29	77.51	192.89	189.91	77.58
16	192.58	105.91	189.80	188.40	180.55	164.42	158.01	188.12	185.89	150.87
17	191.17	106.02	189.90	188.29	180.03	64.20	65.28	187.54	185.26	65.04
18	181.99	106.12	190.01	188.22	180.72	47.38	50.83	188.41	184.49	52.12
19	178.85	106.22	190.11	188.08	182.15	68.99	69.37	188.55	184.63	72.33
20	179.80	106.32	190.21	188.40	179.66	66.42	66.20	187.03	183.19	66.57
21	178.19	106.42	190.32	188.25	177.32	160.71	151.36	186.58	183.38	151.28
22	177.97	106.52	190.42	188.17	177.75	40.43	43.49	188.20	184.19	44.24
23	177.79	106.62	190.53	188.17	176.58	147.66	146.22	189.34	182.98	145.85
24	179.46	106.72	190.63	188.11	178.43	148.57	147.37	189.62	183.31	146.14
25	177.45	106.82	190.74	188.42	177.10	123.66	120.77	186.68	182.11	121.09
26	176.57	106.92	190.84	188.28	176.20	163.36	159.65	185.21	181.32	167.23
27	174.00	107.01	190.95	188.05	174.05	67.27	68.11	181.54	179.22	67.00
28	174.49	107.11	191.05	188.40	173.80	132.51	128.96	180.66	176.58	129.74
29	175.84	107.21	191.16	188.29	177.11	154.12	154.03	182.37	176.96	156.69
30	177.23	107.30	191.26	188.22	177.46	116.20	114.95	183.79	176.15	113.49

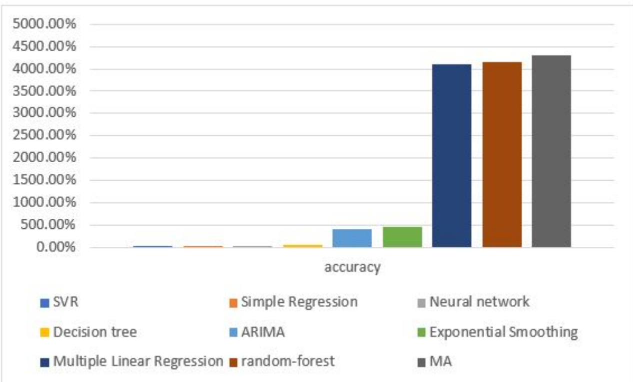

Figure 5: accuracy of forecasting techniques

Table 3: accuracy of forecasting techniques

Methods	accuracy
SVR	1.80 %
Simple Regression	1.97%
Neural network	2.98%
Decision tree	41.34%
ARIMA	4.08%
Exponential Smoothing	4.48%
Multiple Linear Regression	41.09%
random-forest	41.45%
MA	43.13%

straightforward. Neural Networks, a part of deep learning, outperformed the others by spotting complex, non-linear patterns in financial data with ease. They were also flexible in handling lots of data features and automatically finding important patterns. Lastly, ARIMA, designed for time series data, did well at capturing time-related patterns like trends and seasonality. It was easy to understand and had a solid track record in time series predictions. The worse performance belongs to the moving average forecasting method. The moving Average method's poor performance in forecasting financial asset prices is likely due to its simplicity, lack of adaptability, and sensitivity to noise. In financial markets, where data can be complex and influenced by various factors, more advanced methods that can capture both linear and non-linear patterns, adapt to changing conditions, and handle noise effectively is often preferred for accurate forecasting.

5 Conclusion

In this research firstly, the literature about forecasting methods was explained. Then, some of the most used methods in literature were selected by emphasizing the importance of choosing the right forecasting methods in the financial market. These methods included Moving Average (MA), Exponential Smoothing, ARIMA, Simple Linear Regression, Multiple Regression, Decision Trees, Random Forests, Neural Networks, and Support Vector Regression (SVR). These techniques have been widely used in other fields like economic forecasting, quality control, stock market, weather forecasting, etc. The APPLE closing prices in 100 prominent companies dataset, chosen based on revenue, were selected as a case study. the list of 100 big Companies published by The Fortune Global 500. The close price of each company was downloaded from the Yahoo finance library in the period of 4 January 2018 to 22 August 2023 as daily. The forecasting methods were examined by predicting the last 30 days APPLE closing price that was removed from the dataset to evaluate the accuracy of methods based on original and predicted data. The results of this paper show SVR, Simple Linear Regression, Neural Networks, and ARIMA exhibited superior performance in accurately predicting the 30-day APPLE closing prices, while the Moving Average method lagged due to its limitations in handling the complexity of financial data. Malkarjuna and Rao [27], Mondal et al. [30], and Bijesh et al. [12], also emphasized that the ARIMA model is highly accurate in predicting Finance values. Chatzis et al. [10], Ghanbari and Arian [16], and Khan et al. [22] show that the SVR model has good performance in the finance market. In conclusion, the findings of this study emphasize the critical role of selecting the right forecasting method when dealing with financial asset prices. It is important to note that the generalizability of our findings may be constrained to this specific subset of companies, and caution should be exercised when extrapolating these results to a broader range of industries or revenue profiles. SVR, Simple Linear Regression, Neural Networks, and ARIMA are identified as the most effective models for this specific task, showcasing their adaptability, accuracy, and ability to capture both linear and non-linear patterns. Conversely, the Moving Average method's underperformance underscores its unsuitability for this complex task. These insights provide valuable guidance for researchers and practitioners seeking to make informed choices when predicting financial asset prices and highlight the significance of leveraging advanced modeling techniques to navigate the intricacies of financial markets. However, the diversity of industries and financial structures within the Fortune Global 500 list is acknowledged as a potential limitation. Future research could benefit from exploring the performance of these forecasting methods across a more diverse set of companies with varying data properties. By doing so, a more comprehensive understanding of the strengths and weaknesses of each method in different contexts could be achieved.

Bibliography

[1] Armaan, M. S. A., Antony A. S., 2019, A comparison of regression models for prediction of graduate admissions, 2019 International Conference on Computational Intelligence in Data Science (ICCIDS), Chennai, India.

[2] Armstrong, J. S. (Ed.). (2001), *Principles of Forecasting: A Handbook for Researchers and Practitioners* (Vol. 30). Boston, MA: Kluwer Academic.

[3] Banerjee, D., 2014, Forecasting of Indian stock market using time-series ARIMA model. In 2014 2nd international conference on business and information management (ICBIM) (pp. 131-135). IEEE.

[4] Bharathi, S., & Geetha, A. (2017), "Sentiment analysis for effective stock market prediction." In *International Journal of Intelligent Engineering and Systems*, 10(3), 146-154.

[5] Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992), "A training algorithm for optimal margin classifiers." In *Proceedings of the fifth annual workshop on computational learning theory*, 144-152. New York, NY, USA: ACM.

[6] Box, G., 2013, Box and Jenkins: time series analysis, forecasting and control. In A Very British Affair: Six Britons and the Development of Time Series Analysis During the 20th Century, London, UK, Palgrave Macmillan UK.

[7] Breiman, L., 2001, Random forests. Machine Learning 45, 5-32. https://doi.org/10.1023/A:1010933404324

[8] Chandwani, D., & Saluja, M. S. (2014), "Stock direction forecasting techniques: An empirical study combining machine learning system with market indicators in the Indian context." In *International Journal of Computer Applications*, 92(11), 8-17.

[9] Chatfield, C., 2000, Time-series forecasting, CRC Press, 263pp.

[10] Chatzis, S. P., Siakoulis, V., Petropoulos, A., Stavroulakis, E., Vlachogiannakis, N. ,2018, Forecasting stock market crisis events using deep and statistical machine learning techniques. Expert Systems with Applications, 112, 353-371. https://doi.org/10.1016/j.eswa.2018.06.032

[11] Dharmawan, P. A. S., & Indradewi, I. G. A. A. D. (2021), "Double exponential smoothing brown method towards sales forecasting system with a linear and non-stationary data trend." In *Journal of Physics: Conference Series*, IOP Publishing.

[12] Dhyani, B., Kumar, M., Verma, P., Jain, A. ,2020, Stock market forecasting technique using ARIMA model, International Journal of Recent Technology and Engineering, 8(6), 2694-2697. https://doi.org/10.35940/ijrte.f8405.038620

[13] Drucker, H., Burges, C. J. C., Kaufman, L., Smola, A. J., & Vapnik, V. (1997), "Support vector regression machines." In *Advances in Neural Information Processing Systems*, 9, 155-161.

[14] Du, Y., 2018, Application and analysis of forecasting stock price index based on combination of ARIMA model and BP neural network. In 2018 Chinese control and decision conference (CCDC) 2854-2857. IEEE.

[15] Fildes, R., & Kourentzes, N. (2011), "Validation and forecasting accuracy in models of climate change." In *International Journal of Forecasting*, 27(4), 968-995.

[16] Ghanbari, M., Arian, H. ,2019, Forecasting stock market with support vector regression and butterfly optimization algorithm, arXiv preprint arXiv:1905.11462. https://doi.org/10.48550/arXiv.1905.11462

[17] H.-I. Lim, 2019, A Linear Regression Approach to Modeling Software Characteristics for Classifying Similar Software, in 2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC).

[18] Hansun, S., & Subanar, S. (2016). "H-WEMA: A New Approach of Double Exponential Smoothing Method." In *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 14(2), 772-777. http://doi.org/10.12928/telkomnika.v14i2.3096

[19] Hidayatulah, H., & Parasian, S. (2020), "Comparison of forecasting accuracy rate of exponential smoothing method on admission of new students." In *Journal of Critical Review*, 7(2), 268-274.

[20] Hyndman, R. J., & Athanasopoulos, G. (2018), *Forecasting: Principles and Practice*. OTexts. 292 pp.

[21] Hyndman, R. J., Koehler, A. B., Snyder, R. D., & Grose, S. (2002), "A state space framework for automatic forecasting using exponential smoothing methods." In *International Journal of Forecasting*, 18(3), 439-454. https://doi.org/10.1016/S0169-2070(01)00110-8

[22] Khan, W., Ghazanfar, M. A., Azam, M. A., Karami, A., Alyoubi, K. H., Alfakeeh, A. S. ,2020, Stock market prediction using machine learning classifiers and social media, news. Journal of Ambient Intelligence and Humanized Computing, 13, 1-24. https://doi.org/10.1007/s12652-020-01839-w

[23] Khemavuk, P., & Leenatham, A. (2020), "A Conceptual Model for Uncertainty Demand Forecasting by Artificial Neural Network and Adaptive Neuro-Fuzzy Inference System Based on Quantitative and Qualitative Data." In *International Journal of Operations and Quantitative Management*, 26(4), 285-302. http://dx.doi.org/10.46970/2021.26.4.3

[24] Kucharavy, D., Damand, D., & Barth, M. (2023). "Technological forecasting using mixed methods approach." In *International Journal of Production Research*, 61(16), 5411-5435. https://doi.org/10.1016/j.trc.2018.03.001

[25] Kumar, M., Thenmozhi, M. (2014). "Forecasting stock index returns using ARIMA-SVM, ARIMA-ANN, and ARIMA-random forest hybrid models." In *International Journal of Banking, Accounting and Finance*, 5(3), 284-308. https://doi.org/10.1504/IJBAAF.2014.064307

[26] Makridakis, S., Spiliotis, E., Assimakopoulos, V. (2018). "Statistical and Machine Learning forecasting methods: Concerns and ways forward." In *PloS one*, 13(3), e0194889. https://doi.org/10.1371/journal.pone.0194889

[27] Mallikarjuna, M., Rao, R. P. (2019). "Evaluation of forecasting methods from selected stock market returns." In *Financial Innovation*, 5(1), 1-16. https://doi.org/10.1186/s40854-019-0157-x

[28] Meher, B. K., Hawaldar, I. T., Spulbar, C. M., Birau, F. R. (2021). "Forecasting stock market prices using mixed ARIMA model: A case study of Indian pharmaceutical companies." In *Investment Management and Financial Innovations*, 18(1), 42-54. http://dx.doi.org/10.21511/imfi.18(1).2021.04

[29] Meneghini, M., Anzanello, M., Kahmann, A., & Tortorella, G. (2018). "Quantitative demand forecasting adjustment based on qualitative factors: case study at a fast food restaurant." In *Sistemas & Gestão*, 13(1), 68-80.

[30] Mondal, P., Shit, L., Goswami, S. (2014). "Study of effectiveness of time series Modelling (ARIMA) in forecasting stock prices." In *International Journal of Computer Science, Engineering and Applications*, 4(2), 1329. https://doi.org/10.5121/ijcsea.2014.4202

[31] Naik, N., & Mohan, B. R. (2021). "Novel stock crisis prediction techniquea study on Indian stock market." In *IEEE Access*, 9, 86230-86242. https://doi.org/10.1109/ACCESS.2021.3088999

[32] Peng, Z., Li X. (2018). "Application of a multi-factor linear regression model for stock portfolio optimization." In *2018 International Conference on Virtual Reality and Intelligent Systems (ICVRIS).*

[33] Petropoulos, F., Apiletti, D., Assimakopoulos, V., Babai, M. Z., Barrow, D. K., Taieb, S. B., Ziel, F. (2022). "Forecasting: theory and practice." In *International Journal of Forecasting*, 38(3), 705-871. https://doi.org/10.1016/j.ijforecast.2021.11.001

[34] Roopa, H., Asha, T. (2019). "A linear model based on principal component analysis for disease prediction." In *IEEE Access*, 7, pp. 105314-105318, 2019.

[35] Scornet, E., Biau G., Vert, J.-P. (2015). "Consistency of random forests." In *Annals of Statistics*.

[36] Scott, A. J., Fred, C. (2001). *Principles of forecasting: a handbook for researchers and practitioners*. Boston, MA: Kluwer Academic, 862 pp.

[37] Shah, D., Isah, H., & Zulkernine, F. (2019), "Stock market analysis: A review and taxonomy of prediction techniques." In *International Journal of Financial Studies*, 7(2), 26.

[38] Shukor, S. A., Sufahani, S. F., Khalid, K., Abd Wahab, M. H., Idrus, S. Z. S., Ahmad, A., & Subramaniam, T. S. (2021, May). "Forecasting Stock Market Price of Gold, Silver, Crude Oil, and Platinum by Using Double Exponential Smoothing, Holts Linear Trend, and Random Walk." In *Journal of Physics: Conference Series (Vol. 1874, No. 1, p. 012087)*. IOP Publishing. https://doi.org/10.1088/1742-6596/1874/1/012087

[39] Siami-Namini, S., Tavakoli, N., Namin, A. S. (2018). "A comparison of ARIMA and LSTM in forecasting time series." In *17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pp. 1394-1401. https://doi.org/10.1109/ICMLA.2018.00229

[40] Sidqi, F., & Sumitra, I. D. (2019). "Forecasting product selling using single exponential smoothing and double exponential smoothing methods." In *IOP conference series: materials science and engineering*.

[41] 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." In *International Journal of Financial Studies*, 11(3), 94. https://doi.org/10.3390/ijfs11030094

[42] Taylor, J. W. (2008). "An evaluation of methods for very short-term time series forecasting." In *International Journal of Forecasting*, 24(4), 635-642. https://doi.org/10.1016/j.ijforecast.2008.07.007

[43] Uyank, G. K., Güler, N. (2013). "A study on multiple linear regression analysis." In *Procedia-Social and Behavioral Sciences*, 106, 234-240. https://doi.org/10.1016/j.sbspro.2013.12.027

[44] Voyant, C., Notton, G., Kalogirou, S., Nivet, M. L., Paoli, C., Motte, F., Fouilloy, A. (2017). "Machine learning methods for solar radiation forecasting: A review." In *Renewable energy*, 105, 569-582. https://doi.org/10.1016/j.renene.2016.12.095

[45] Wager, S., & Athey, S. (2018). "Estimation and inference of heterogeneous treatment effects using random forests." In *Journal of the American Statistical Association*, 113, 1228-1242. https://doi.org/10.48550/arXiv.1510.04342

[46] Wang, X., Sun X. (2016). "An improved weighted naive Bayesian classification algorithm based on multivariable linear regression model." In *2016 9th International Symposium on Computational Intelligence and Design (ISCID).*

[47] Wu, J., Liu, C., Cui W., & Zhang Y. (2019). "Personalized Collaborative Filtering Recommendation Algorithm based on Linear Regression." In *2019 IEEE International Conference on Power Data Science (ICPDS).*

[48] Wu, Y., Tan, H., Qin, L., Ran, B., & Jiang, Z. (2018). "A hybrid deep learning based traffic flow prediction method and its understanding." In *Transportation Research Part C: Emerging Technologies*, 90, 166-180.

[49] Zeroual, A., Harrou, F., Dairi, A., & Sun, Y. (2020). "Deep learning methods for forecasting COVID-19 time-series data: A Comparative study." In *Chaos, solitons & fractals*, 140, 110121. https://doi.org/10.1016/j.chaos.2020.110121

[50] Zhang, F., & O'Donnell, L. J. (2020). "Support vector regression." In *Machine learning*, 123-140, Academic Press. https://doi.org/10.1016/B978-0-12-815739-8.00007-9

How to Cite: Mahdi Goldani[1], *Comparative analysis on forecasting methods and how to choose a suitable one: case study in financial time series*, Journal of Mathematics and Modeling in Finance (JMMF), Vol. 3, No. 2, Pages:35–59, (2023).