International Journal of Innovative Research in Computer Science and Technology (IJIRCST)
ISSN (online): 2347-5552, Volume-11, Issue-6, November 2023
https://doi.org/10.55524/ijircst.2023.11.6.4
Article ID IRP1414, Pages 18-25
www.ijircst.org

Applications of Machine Learning in Predictive Analysis and Risk Management in Trading

Kavin Karthik V

B.Tech Scholar, School of Computer Science and Engineering (SCOPE), Vellore Institute of Technology, Chennai, Tamil Nadu, India

Correspondence should be addressed to Kavin Karthik V; kavinkarthivs@gmail.com

Received 23 October 2023; Revised 7 November 2023; Accepted 20 November 2023

ABSTRACT- The stock market is considered the primary domain of importance in the financial sector where Artificial Intelligence combined with various algorithmic practices empowers investors with data-driven insights, enhancing decision-making, predicting trends, and optimizing risk management for more informed and strategic financial outcomes. This research paper delves into the real-world applications of machine learning and algorithmic trading, observing their historical evolution together and how both of these can go hand in hand to control risk and forecast the movement of a stock or an index and its future. The research is structured to provide comprehensive insights into two major subdomains in the application of AI in algorithmic trading: risk management in equity markets and predictive analysis of stock trends through the application of machine learning models and training the current existing data which is feasible and training them with respect to historical scenarios of various market trends along with various fundamental and technical analysis techniques with the help of various deep learning algorithms. For risk management of a portfolio in finance, various machine learning models can be employed, depending on the specific needs and goals of the portfolio manager or risk analyst and implementing various value-at-risk algorithms along with deep learning techniques in order to assess risk at particular trade position and to manage volatile trades at unprecedented situations. The significance of this research paper lies in its practical applicability, offering real-world solutions to enhance trading strategies and decision-making processes with a focus on mitigating risk and capitalizing on market opportunities and also giving clear insights with respect to the current practical limitations of application of the provided solution and future scope to overcome the same.

KEYWORDS- Algorithmic Trading, Risk Management, Equity Markets, Portfolio Management, Predictive Analysis, Fundamental Analysis, Value at Risk.

I. INTRODUCTION

The intersection of artificial intelligence (AI) and financial trading has evolved into an intricate symbiosis, catalyzing the transformation of traditional trading practices. This research paper embarks on an in-depth exploration of this dynamic relationship, dissecting its historical progression and unveiling the profound implications for modern trading strategies. The paper is structured to provide a comprehensive understanding of two pivotal facets: risk management in equity markets and predictive analysis of stock trends through the application of cutting-edge machine learning models, including Convolutional Neural Networks (CNN), Long Short-Term Memory networks (LSTM), and Deep Reinforcement Learning (DRL).

The historical roots of AI in trading trace back to the mid-20th century when early AI pioneers, like Marvin Minsky and John McCarthy, envisaged the potential of machines to replicate human decision-making in complex financial domains. In recent decades, advancements in computational power, access to extensive historical financial data, and the development of sophisticated machine learning algorithms have propelled the integration of AI in trading systems. This evolution has given rise to algorithmic trading, high-frequency trading, and quantitative strategies, altering the landscape of financial markets [1].

Risk management and portfolio optimization have experienced a transformative historical journey. Traditionally, risk management relied on statistical methods and historical data to estimate portfolio risk. Markowitz's Mean-Variance Optimization (MVO), introduced in 1952, revolutionized portfolio optimization. MVO aimed to strike an optimal balance between risk and return by considering asset correlations. However, it had limitations in handling non-linear relationships and outliers, which are prevalent in financial markets. The historical development of AI techniques in portfolio optimization can be traced back to the late 20th century[4]. Neural networks emerged as promising tools for modeling complex financial data. Multilayer perceptrons, a type of neural network, could capture non-linear patterns and improve risk estimation. Nevertheless, their training process required extensive data, making them less suitable for small datasets. Convolutional Neural Networks (CNNs), initially designed for image analysis, found their way into financial markets. CNNs can process time series data and identify complex patterns, which is pivotal for portfolio optimization.[9] For example, they can analyze historical price charts and uncover non-linear relationships between assets, enhancing risk estimation.

The way of finding the future valuation of the stock market prices is called the stock market estimate.

Expected to be Strong, accurate, and effective. The system should work in line with real-life scenarios and be well connected to that. The movement in the stock market is usually determined by the sentiments of thousands of investors [2]. These events are political events such as the statements of ministers or government officials, statements of government bodies such as RBI, SEBI, scandal news, etc. It can also be the global happening such as rapid movements in currencies prices and commodities prices [3]. All this thing affects the earnings of companies, which ultimately affects the sentiment of stock market investors. This is beyond the reach of almost all individuals to assess these accurately and consistently. This method usually requires the collection of various social media data, news that affects stock market investors sentiment, and the feelings expressed by individuals. Other data such as last year's stock prices are also considered. The relationship between different data points. Is considered and an estimation is done using these variety of data points. Robust risk management in equity markets is a cornerstone of successful trading. AI contributes significantly by enabling the development of advanced risk models, including Value at Risk (VaR) and Conditional Value at Risk (CVaR) [2]. These models employ machine learning and statistical techniques to assess market volatility, optimize portfolio diversification, and formulate effective hedging strategies [4]. Predictive analysis of stock trends is a fundamental concern for traders and investors. CNN, a neural network architecture designed for image processing, can be adapted to identify patterns in financial time series data. [5] LSTM, a specialized recurrent neural network, excels in modeling sequential data and is well-suited for predicting stock price movements. DRL, inspired by reinforcement learning, offers a dynamic approach to trading by learning optimal strategies through interactions with the market environment [6].

Convolutional Neural Networks (CNN) in Stock Price Prediction: CNN, originally developed for image processing, has found a unique application in financial markets. These networks can analyze historical price charts, identifying patterns and trends that may not be apparent to human analysts. By processing the pixel-level data of price charts, CNN models can extract valuable information to predict short-term and long-term stock price movements. Their ability to recognize complex patterns in financial time series data makes them a valuable tool for traders and investors [7]. Long Short-Term Memory Networks (LSTM) in Time Series Analysis: LSTM, a specialized recurrent neural network architecture, has gained immense popularity in time series forecasting. These networks are well-suited for modeling sequential data, which is a key characteristic of financial time series. LSTM models can capture dependencies over time, making them effective for predicting stock prices and trends. Their ability to remember and learn from past data enables them to adapt to changing market conditions and provide accurate forecasts [8]. Deep Reinforcement Learning (DRL) for Adaptive Trading: Deep Reinforcement Learning, inspired by reinforcement learning principles, offers a dynamic approach to trading. DRL agents learn optimal strategies by interacting with financial markets, receiving rewards for profitable actions and penalties for losses[9]. Over time, these agents develop adaptive trading strategies that can capitalize on changing market conditions. This adaptability makes DRL a valuable tool for traders seeking to navigate the complexities of financial markets and make decisions in real time based on the evolving market landscape.

II. LITERATURE SURVEY

The paper is divided into 5 major sections: Historical evolution and the current trend of AI and Algorithmic Trading, Previous research and articles in the same domain, Methodology of assessing risk in executing a trade, Methodology to forecast a trade, Real-world application of the methodology and algorithm, Limitations in the methodologies and Scope for future improvement. My contribution to this domain of research majorly involves Long Short-Term Memory (LSTM) and deep learning algorithms in the context of stock market predictive analysis and making use of Collaborative Filtering and value-at-risk methodologies for risk management in portfolio optimization.

Nayak et al [17] highlight precision and accuracy in predicting share prices. While various methods, including time series analysis, fundamental analysis, and technical analysis, have been employed by investors and institutions, they are not always reliable. To address this challenge, the paper introduces a Machine Learning (ML) approach, which is trained on historical stock data to acquire intelligence and make accurate predictions. After extensive research, the paper finds that the Artificial Neural Network (ANN) is more suitable than other algorithms. A customized neural network model is employed, and the method is tested on the Bombay Stock Exchange index dataset

Ding, G., et al. [18] explores a two-layer approach incorporating technical analysis and machine learning. The model described in the paper incorporates a two-layer approach, with the first layer relying on technical analysis and the second on machine learning. Additionally, the paper introduces a financial management strategy that considers the historical success of predictions to determine future investments. Through portfolio simulations and trading models, the research concludes that the predictive model effectively surpasses the Oslo Benchmark Index (OSEBX) [7]

Birant, D et al [19] proposes a model combining particle swarm optimization and random forest algorithms for daily share market price prediction.[9] This model works with historical share market data and technical indicators. By avoiding local minima and improving prediction accuracy, the particle swarm optimization algorithm plays a key role. The model is tested on multiple financial datasets and compared with Leuralberg-Marquardt neural network algorithms, resulting in improved prediction accuracy.

A. Nayak, M. M. M. Pai, and R. M. Pai's "Prediction Models for Indian Stock Market" [17] discusses predictive models for next-day and one-month share price predictions using supervised machine learning. Daily predictive models rely on historical data, and machine learning techniques yield up to 70% accuracy. Monthly forecasting models attempt to identify patterns between different months. Tests indicate that the trend for at least one month is correlated with other months' trends,

reinforcing the importance of machine learning in predicting share market prices.
G. Ding and L. Qin's "Study on the Prediction of Stock Price Based on Associated Network Model of LSTM" [18] explores multiple methods to estimate future share prices, including various AI techniques. The paper introduces a deep neural network model based on Long Short-Term Memory (LSTM) networks with multiple inputs and outputs. This model can simultaneously estimate open, high, and low share prices. The results demonstrate that the associated model surpasses other models in accuracy, achieving an estimated accuracy of over 95%.
S. P. Pimpalkar, Jenish Karia, Muskaan Khan, Satyamandand, and Tushar Mukherjee's "Stock Market Forecasts Using Machine Learning" [22] explores the construction of predictive models using various attributes, including oil prices, foreign exchange rates, interest rates, gold and silver prices, news, Twitter news feeds, and pattern matching. The paper employs different Machine Learning (ML) methods, such as Support Vector Machines and Recurrent Neural Networks, to predict market movements based on these attributes.
M. Moukalled, W. El-Hajj, and M. Jaber's "Automated Stock Price Prediction Using Machine Learning" [20] highlight the importance of news in share market predictions. The authors employ conventional machine learning models and deep learning models to predict stock prices by considering relevant issues. The research reveals that a Support Vector Machine model achieves the highest accuracy of 82.91% for predicting stock prices.
V. V. K. Sai Reddy's "Stock Market Forecasts Using Machine Learning" [21] emphasizes the importance of stock market prediction, employing Machine Learning (ML) methods to predict major and minor stock prices in different markets. The paper introduces a Support Vector Machine for predicting share prices using daily and minute wave prices [16].
When it comes to Risk Management, Collaborative filtering (CF) is one of the most popular techniques used to build recommender systems. It works on the assumption that users with similar preferences in the past are more likely to have similar interests in the future. There are two categories: User-based and Item-based. User-based CF finds users with similar consumption patterns and recommends items that these similar users find interesting. User-based CF performance decreases when we have sparse datasets (Boström & Filipsson, 2017). On the other hand, item-based CF recommends items similar to what the user has liked in the past. Since the properties of items remain more constant compared to users' evolving tastes and preferences, item-based CF tends to be more scalable than user-based CF (Boström & Filipsson, 2017).
Sayyed et al. (2013) [22] presented a preliminary investigation on the implementation of CF methods on the stock market but had not validated CF methods on stock data. Vismayaa et al. (2020) compared the performance of single classifiers against ensemble classifier-based recommender systems on stocks from the Bombay Stock Exchange based on classification accuracy and economic measures. This study showed that ensemble classifier-based recommender systems surpass single classifiers in predicting stock price movements.
Prieto-Torres and Galpin [24] used a CF model to build a virtual wallet recommender system for a FinTech company to enhance the user experience on their mobile application. However, this method worked best using a non-personalized approach where the top-N most popular stocks were recommended to each user. Anwar, et.al, (2019) proposed a collaborative filtering-based movie recommendation using rule mining to satisfy the individual requirement. Their paper also provided a comparative analysis of the several types of similarity measures namely Co- sine, Correlation, Euclidean, Jaccard, and Manhattan. The correlation similarity measure yielded the best precision, recall, and F1-score. Further, Anwar, et.al, (2021) introduced a new approach that uses CF and Singular Value Decomposition (SVD) ++ for implementing a recommendation system. Their proposed approach gave a smaller error rate when cross-validation (CV = {5, 10, 15}) was performed. Their proposed approach also alleviated the sparsity and cold-start problems.
Kirubahari, R, et.al [23] proposed a weighted parallel deep hybrid collaborative filtering approach based on Singular Value Decomposition (SVD) and Restricted Boltzmann Machine (RBM). Their results indicated better prediction compared to other techniques in terms of ac-curacy. Cui et. al (2019) also explored the traditional CF approach as well as SVD to ease fund companies' decision-making process in identifying which listed company to invest in. In their study, they tackled the issue of sparsity for user-based CF by taking into account the number of items that are similarly scored by the users and improved traditional item-based CF by inputting the missing ratings of the users through an iterative process instead of just using the mean ratings. Their study showed that the improved item-based CF outperforms the other two methods.
Lastly, Rukiya et. al [25] proposes an improvement to the traditional CF approach by first clustering based on the popularity of the items and segmenting the users based on their loyalty to the business to provide a more personalized recommendation while ensuring that items that are already known to the user are not recommended. This allows e-commerce platforms' targeted marketing approaches to be more effective, garnering greater user loyalty. Content-based filtering (CBF) is another technique used to build recommender systems. It uses the information about the items obtained from the user's profile to recommend similar items to what the user has liked in the past. [12] Unlike CF, CBF does not take into account other users' information. Yoo et al. (2003) proposed a personalized filtering system that only considers the trading information that a user deems relevant. Personalized recommendations based on the Moving Average Convergence Divergence (MACD) advise traders whether to buy or sell a stock by examining the difference between two moving averages. Chalida bhongse and Kaensar (2006) proposed an adaptive user model that performs stochastic technical analysis to forecast stock returns while considering explicit preferences and user interactions with the system for personalized recommendations. Rutkowski (2021) built a Neuro-Fuzzy recommender system to recommend a stock

based on similarity from a user's past transaction if the stock passed the fuzzy set of rules generated.

III. METHODOLOGY

SMP systems can be classified according to the type of data they use as the input. Most of the studies used market data for their analysis. Recent studies have considered textual data from online sources as well. In this section, the studies are classified based on the type of data they use for prediction purposes.

Market Data Market data are the temporal historical price-related numerical data of financial markets. Analysts and traders use the data to analyze the historical trend and the latest stock prices in the market. They reflect the information needed for the understanding of market behavior. The market data are usually free, and can be directly downloaded from the market websites. Various researchers have used this data for the prediction of price movements using machine learning algorithms. Previous studies have focused on two types of predictions. Some studies have used stock index predictions like the Dow Jones Industrial Average (DJIA) [7], Nifty [9], Standard and Poor's (S&P) 500 [10], National Association of Securities Dealers Automated Quotations (NASDAQ) [12], the Deutscher Aktien Index (DAX) index [13], and multiple indices [14,15]. Other studies have used individual stock prediction based on some specific companies like Apple, Google, or groups of companies. Furthermore, the studies focused on time-specific predictions like intraday, daily, weekly, and monthly predictions, and so on. Moreover, most of the previous research is based on categorical prediction, where predictions are categorized into discrete classes like up, down, positive, or negative [13]. Technical indicators have been widely used for SMP due to their summative representation of trends in time series data. Some studies considered different types of technical indicators, e.g., trend indicators, momentum indicators, volatility indicators, and volume indicators. Furthermore, numerous studies have used an amalgam of different types of technical indicators for SMP.

Textual Data Textual data is used to analyze the effect of sentiments on the stock market. Public sentiments have been proven to affect the market considerably. The most challenging part is to convert the textual information into numerical values so that it can be fed to a prediction model. Furthermore, the extraction of textual data is a challenging task. The textual data has many sources, such as financial news websites, general news, and social platforms. Most of the studies were carried out on textual data to try to predict whether the sentiment towards a particular stock is positive or negative. The previous studies considered several textual sources for SMP, such as the Wall Street Journal, Bloomberg, CNBC and Reuters, Google Finance, and Yahoo Finance. The extracted news may be either generalized news or some specific financial news, but the majority of the researchers use financial news, as it is deemed to be less susceptible to noise. Some researchers have used less formal textual data, such as message boards. Meanwhile, the textual data from microblogging websites and social networking websites are comparably less explored than other textual data forms for SMP. Besides this, one challenge faced for the processing of the textual data are that the information generated on these platforms is enormous, increasing the computational complexities. For example, the researchers in [13] processed 1,00,000 tweets, and the researchers in [12] processed around 2,500,000 tweets, which was a complex task. Moreover, for the textual data, no proper standard format is followed while posting on social media, which increases the processing complexities. In addition, the detection of shorthand spellings, emoticons, and sarcastic statements are yet another challenge. Machine learning algorithms come to the forefront to deal with all kinds of challenges faced while processing textual data.

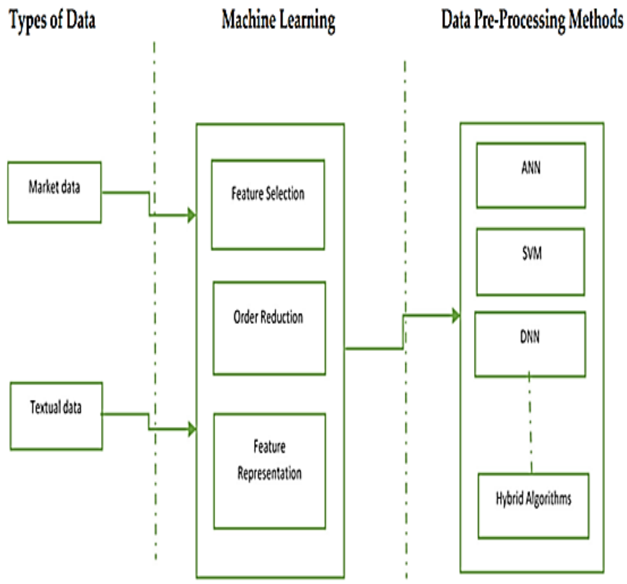

Figure 1: Generic Scheme for Stock Market Prediction using ML Algorithms

IV. DATA PREPROCESSING

Once the data is available, it needs some pre-processing so that it can be fed to a machine-learning model. The significance of the output depends on the pre-processing of the data [15]. The textual data must be transformed into a structured format that can be used in a machine-learning model. The previous studies revealed that there are three significant pre-processing steps, i.e., feature selection, order reduction, and the representation of features. Table 1 presents the comparison of the data sources, type of input, and prediction duration. Table 2 presents the comparison of the data pre-processing techniques used in the studies so far and how each of the techniques contributes to enhancing accuracy in each method.

Table 1: Comparison of the data sources, type of input, and prediction

Data	Type of Input	Prediction Duration
S&P 500	Market data	Few days ahead
NASDAQ index	Market data	Few days ahead
DAX 30	broker house newsletters, RSS market feeds, and stock exchange data	Intraday
Yahoo Finance	Financial News	Intraday
DGAP, Euro-Adhoc	Corporate announcements financial new	Daily
Yahoo finance (18 Stock Companies data)	Market data, yahoo finance message board data	Daily
DJIA	Market data and Twitter	Daily
BSE and NSE stocks	Market data, technical indicators, Twitter data	Intraday
Nifty and Sensex	Market data and news	Intraday

Table 2: Comparison of the data pre-processing techniques

Feature Selection	Order Reduction	Feature Representation
Bag of words, LDA, JST, Aspect Based	-	TF-IDF
Correlation	Lemmatization	Boolean
Bag of Words	Chi2, Information Gain, Document Frequency, Occurrence	TF-IDF
Bag-of-word, Word2vec		TF-IDF
GA	PCA, FA, FO	-
N-grams	SVM based Recursive Feature Elimination, PCA, KPCA, and XGB	-
Bag-of-words	Occurrence	TF-IDF
GA, Feature Ranking	PCA-SVM, DA-RNN	-

V. PROPOSED SYSTEM

In the proposed system, our primary objective is to predict future share prices using various advanced machine learning methods, with a specific focus on deep neural networks like LSTM (Long Short-Term Memory). Our approach involves training and testing machine learning algorithms on historical data points, allowing us to estimate future share prices with accuracy and precision [15]. To accomplish this, we leveraged a rich dataset comprising EOD (End of the Day) historical data spanning several years. This dataset serves as the foundation for training and testing our machine-learning models [15]. We employed a range of machine learning libraries and frameworks to achieve our goals, including NumPy and Pandas for data manipulation and visualization, Scikit-Learn (Sklearn) and TensorFlow for building machine learning models, and the math library for performing mathematical operations on the data [6].

A significant component of our dataset for sentiment analysis was historical news headlines. This data source contributes to a deeper understanding of market sentiment and enhances the prediction accuracy.

In our data processing pipeline, the NumPy library played a pivotal role in preparing and cleaning the datasets. It allowed us to transform the data into a format that can be directly used by the machine learning models [6].

The Sklearn library is essential for actual calculations, estimations, and predictions. It serves as a reliable tool for performing various machine-learning tasks, ensuring that our predictive models are robust and accurate [4]. We sourced historical share market data from various public and open online repositories, using approximately 85% of the dataset for training the machine learning models, with the remainder allocated for validation and testing purposes [4].

The core approach of our supervised machine learning models is to analyze and recognize patterns and correlations within the dataset. By training on both the training and validation datasets, the models learn to produce accurate predictions for the test dataset [13].Our system undertakes feature extraction and various data pre-processing steps to prepare the data for model training. The Python library, Pandas, played a vital role in combining various datasets into a single, comprehensive data frame. This pre-processed data frame includes critical features such as date, close, open, high, low, volume, delivery percentage, number of trades, turnover, and other derived features. All these features are used to train the machine learning models, with a particular focus on a random forest model. Additionally, we utilize selected features to make predictions using advanced algorithms like LSTM and SVM for forecasting the share prices of the upcoming day [14]. The accuracy of these predictions is thoroughly assessed on the test dataset and compared against real values. Furthermore, our system incorporates Artificial Neural Networks (ANNs) for sentiment analysis of news data. We employ the Google News API to fetch real-time news data related to the stock market. The analysis of this news data is crucial in understanding the market sentiment, which is a key factor in predicting share prices. In terms of risk management, our system integrates risk management algorithms to ensure that investments are well-balanced and aligned with the risk tolerance of the investor. These algorithms analyze portfolio composition and dynamically adjust it based on real-time market conditions, further enhancing the robustness of our system.

Overall, our proposed system encompasses various technical aspects, including data manipulation, feature extraction, sentiment analysis using ANNs, advanced machine learning techniques like LSTM, SVM, and risk management algorithms, all aimed at providing accurate and reliable predictions for share prices in the dynamic stock market environment.

VI. RESULTS AND DISCUSSION

In this section, the results after carrying out the successful implementation of the project have been presented in the form of tables and graphs. As part of this project, five algorithms i.e., the K-Nearest Neighbour algorithm, Support Vector Regression algorithm, Linear Regression algorithm, Decision Tree Regression algorithm, and Long short-term memory algorithm were chosen for the prediction of stock prices of twelve different companies. The dataset was huge, starting from 2015 up to and including 2021. The models were tested for 8 trading days i.e., since the train-to-test ratio was 99:1, and the data for 2304 days was used for training whereas the remaining 8 days were allocated for testing the created models. The models were tested on three essential performance metrics, namely, Symmetric Mean Absolute Percentage Error (SMAPE), R-squared value (R2), and Root Mean Square Error (RMSE). These are well-known prevalent evaluation parameters that help researchers to draw conclusions about the different models that were being studied.

Table 3: Results for the R-Squared value (R2)

Parameter	R^2 (R squared)				
Algorithms	**K-Nearest Neighbors**	**Linear Regression**	**Support Vector Regression**	**Decision Tree Regression**	**Long Short-Term Memory**
Adani Ports	-0.22	-6.67	-1.76	-3.25	-0.90
Asian Paints	-1.66	-5.57	-2.21	-2.75	-0.45
Axis Bank	-5.32	-3.43	-1.05	-4.18	0.59
HDFC	-1.56	-2.47	-2.77	-3.04	-0.62
Hindustan Unilever	-2.03	-1.07	-0.35	-2.39	0.27
ICICI Bank	-2.59	-1.59	-2.72	-2.23	0.45
Kotak Bank	-3.78	-2.31	-3.38	-3.90	-0.01
Maruti	-1.01	-0.65	-2.61	-0.73	-0.99
NTPC	-6.90	-2.31	-1.11	-0.84	-0.02
Tata Steel	-2.16	0.48	-1.13	-2.01	0.80
TCS	-0.83	-1.18	-0.51	-0.92	-0.84
Titan	-0.21	-0.57	-0.12	-1.72	0.31
Average	**-2.42**	**-2.27**	**-1.69**	**-2.33**	**-0.11**

Table 3 displays the tabulated results for the R-Squared value (R2) acquired when a particular model was tested for that particular company's dataset. With the ideal value for R-squared being close to a non-negative '1', from this table, it is observed that out of all the five different algorithmic models, the DL algorithm i.e., the Long Short-Term Memory algorithm has provided the best results, as it has R-squared quite close 1, (i.e., -0.11), followed by Support Vector Regression, with an R-squared value of -1.69, so on and so forth.

Table 4: Results for the SMAPE acquired

Parameter	SMAPE (Symmetric Mean Absolute Percentage Error)				
Algorithms	**K-Nearest Neighbors**	**Linear Regression**	**Support Vector Regression**	**Decision Tree Regression**	**Long Short-Term Memory**
Adani Ports	10.99	9.18	9.51	10.68	1.65
Asian Paints	11.63	9.35	8.42	9.78	1.67
Axis Bank	16.67	10.37	5.64	8.48	1.88
HDFC	20.46	11.00	6.22	12.56	2.19
Hindustan Unilever	10.95	9.62	4.88	8.82	1.38
ICICI Bank	14.45	8.92	7.02	7.37	2.31
Kotak Bank	12.44	10.51	5.81	10.26	1.43
Maruti	15.92	11.13	3.09	13.92	1.32
NTPC	13.59	12.39	3.08	9.60	1.13
Tata Steel	16.08	13.10	5.75	8.06	1.75
TCS	15.84	9.95	4.41	10.05	1.40
Titan	12.90	3.50	3.33	11.39	1.06
Average	**14.32**	**9.91**	**5.59**	**10.08**	**1.59**

Table 4 displays the tabulated results for the SMAPE acquired when a particular model was tested for that particular company's dataset. With the ideal value for SMAPE being close to zero, from this table, it is observed that out of all the five different algorithmic models, the DL algorithm i.e., the Long Short-Term Memory algorithm has rendered the best predictive performance, as it has the least value of error (1.59), followed by Support Vector Regression, with a SMAPE of 5.59 etc.

VII. FUTURE SCOPE AND LIMITATIONS OF PROPOSED MODEL

With the increasing demand for ML in almost every possible place and situation, be it industries business models, or healthcare domains, it is of utmost importance to make better models that can make more accurate and precise predictions from huge sets of data available.

However, the analysis from the results above and various other literature reviews suggests that ML yields less authentic results when it comes to the prediction of time series data. Nevertheless, there is a potent solution to this which lies in DL and neural networks that offer great results during time series prediction when compared to ML. The results found after the implementation of this project conform to the theoretical facts related to the performance of the algorithms. As far as this project is concerned, there may be more algorithms implemented in the future and different datasets may be used in order to offer a wider spectrum for comparison of the algorithms and on a wider note comparison of ML and DL. DL techniques comprise several layers which drive them a step closer to the way human brains function whereas, on the contrary, ML still requires a lot more human assistance. DL techniques are based on the concept of neural networks, similar to the neuron cells in the human brain. Further DL techniques equip the machine to identify features that make identification easier or in other words hierarchical arrangement of the features. A major advantage that comes with DL is that they continue to become more robust and efficient with increasing data, which is a quite favorable feature for data-extensive projects and historic time-series predictions. The future lies in the development and betterment of more and more ML and DL-based models requiring minimum human intervention, lesser prediction time, more accurate predictions, the capability to handle huge data with the help of multiple layers, reduced complexity, and more affordability, etc. The upcoming researchers possess DL and ML possess high potential for the upcoming researchers to study, scientifically curate, and develop much more efficient and robust models that could predict faster and more precise results based on real-time situations and would help the world save and progress in multiple domains to give rise to holistic development in interdisciplinary areas with minimum to no human intervention

This research work aimed at developing ML models that would be capable of predicting stock prices with increased accuracy so that interested traders and investors could make use of such methods to experience increased profits by investing on the right day at the right place. Successful implementation as part of this project was carried out wherein five algorithms i.e., Long Short-Term Memory and deep Neural Network algorithms were deployed to create precise predictive models for application in stock price prediction of twelve prevalent Indian Companies namely, Adani Ports, Asian Paints, Axis Bank, Housing Development Finance Corporation Limited (HDFC) Bank, Industrial Credit and Investment Corporation of India (ICICI) Bank, Kotak Bank, Hindustan Unilever Limited, Maruti, National Thermal Power plant Corporation (NTPC), Tata Steel, Tata Consultancy Services (TCS) and Titan, after which an elaborate comparative analysis of the performances of the algorithms during stock price prediction has been carried out. The stock prices collected were from 2015 to 2021, and after this exhaustive research, it can be concluded that DL algorithms have a substantial edge over simple ML algorithms when it comes to the prediction of time series data. out of the five chosen algorithms, the Long Short-Term Memory algorithm was a DL algorithm that provided the best results during stock price prediction.

REFERENCES

1. Russell, S. J., & Norvig, P. (2016). Artificial Intelligence: A Modern Approach (3rd ed.). Pearson.
2. Chan, E. P. (2013). Algorithmic Trading: Winning Strategies and Their Rationale. John Wiley & Sons.
3. Ashish Sharma, Dinesh Bhuriya, Upendra Singh. "Survey of Stock Market Prediction Using Machine Learning Approach", ICECA 2017.
4. McNeil, A. J., Frey, R., & Embrechts, P. (2015). Quantitative Risk Management: Concepts, Techniques, and Tools (Revised ed.). Princeton University Press.
5. Zhang, Y., Zhao, D., & Dong, F. (2018). Stock Price Prediction with Convolutional Neural Network. In Proceedings of the International Conference on Neural Information Processing (ICONIP) (pp. 843-850).
6. Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation, 9(8), 1735-1780.
7. Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., & Riedmiller, M. (2015). Human-level Control through Deep Reinforcement Learning. Nature, 518(7540), 529-533.
8. Zhang, Y., Zhao, D., & Dong, F. (2018). Stock Price Prediction with Convolutional Neural Network. In Proceedings of the International Conference on Neural Information Processing (ICONIP) (pp. 843-850).
9. Park, S., Kihara, K., & Watanabe, K. (2015). Portfolio Risk Assessment Using Support Vector Machines. In Proceedings of the International Conference on Neural Information Processing (ICONIP) (pp. 226-233).
10. Bansal, M., Goyal, A., & Choudhary, A. (2022). Stock market prediction with high accuracy using machine learning techniques. Procedia Computer Science, 215, 247-265.
11. Rouf, N., Malik, M. B., Arif, T., Sharma, S., Singh, S., Aich, S., & Kim, H. C. (2021). Stock market prediction using machine learning techniques: a decade survey on methodologies, recent developments, and future directions. Electronics, 10(21), 2717.
12. Chandrinos, S. K., Sakkas, G., & Lagaros, N. D. (2018). AIRMS: A risk management tool using machine learning. Expert Systems with Applications, 105, 34-48.
13. Grote, M., & Bogner, J. (2023). A Case Study on AI Engineering Practices: Developing an Autonomous Stock Trading System. arXiv preprint arXiv:2303.13216.
14. El Hajj, M., & Hammoud, J. (2023). Unveiling the Influence of Artificial Intelligence and Machine Learning on Financial Markets: A Comprehensive Analysis of AI Applications in Trading, Risk Management, and Financial Operations. Journal of Risk and Financial Management, 16(10), 434.
15. Mintarya, L. N., Halim, J. N., Angie, C., Achmad, S., & Kurniawan, A. (2023). Machine learning approaches in stock market prediction: a systematic literature review. Procedia Computer Science, 216, 96-102.
16. Hegazy, O., Soliman, O. S., & Salam, M. A. (2014). A machine learning model for stock market prediction. arXiv preprint arXiv:1402.7351.
17. Nayak, A., Pai, M. M., & Pai, R. M. (2016). Prediction models for Indian stock market. Procedia Computer Science, 89, 441-449.
18. Ding, G., & Qin, L. (2020). Study on the prediction of stock price based on the associated network model of LSTM. International Journal of Machine Learning and Cybernetics, 11, 1307-1317.
19. Birant, D., & Işık, Z. (2019, October). Stock market forecasting using machine learning models. In 2019

Innovations in Intelligent Systems and Applications Conference (ASYU) (pp. 1-6). IEEE.

20. Moukalled, M. I. (2019). Automated stock price prediction using machine learning (Doctoral dissertation).
21. Vanukuru, Kranthi. (2018). Stock Market Prediction Using Machine Learning. 10.13140/RG.2.2.12300.77448.
22. F.R. Sayyed, R.V. Argiddi, S.S. Apte. (2013). Collaborative Filtering Recommender System for Financial Market, International Journal of Engineering and Advanced Technology (IJEAT).
23. R, Kirubahari & Joe Amali, Miruna. (2021). A Hybrid Deep Collaborative Filtering Approach for Recommender Systems. 10.21203/rs.3.rs-651522/v1.
24. Prieto-Torres D.R., Galpin I. (2020). A Virtual Wallet Product Recommender System Based on Collaborative Filtering. In: Florez H., Misra S. (eds) Applied Informatics. ICAI 2020. Communications in Computer and Information Science, vol. 1277. Springer, Cham.
25. D. Rukiya, T. Africa, S. Nayak, S. Nooreain. (2019). Recommendation in E-Commerce using Collaborative Filtering International Research Journal of Engineering and Technology, 6 (5)