STOCK MARKET PREDICTION USING MACHINE LEARNING AND DEEP LEARNING TECHNIQUES

Prerana C[1], Pratheeksha Mahishi J[1], N Tahmin Taj[1], Anusha B Shetty[1], Madhu B R[2]

[1]UG Student, [2]Dept. of Computer Science & Engineering, Jyothy Institute of Technology, Bangalore, India

---***---

Abstract – *Prediction of the stock market value has always been an area of interest for investors and researchers for a long time due to its complexity, intrinsic volatility and regularly changing nature. Thus, making reliable predictions are challenging. Both stock prices and twitter data have been employed in the prediction process. Previously various machine learning algorithms have been implemented to predict stock market variations. The objective of this is to compare the methods implemented in the past for the prediction of stock market value. The comparison between the results of the previous work help in determining the efficient algorithm and building the system model based on that selected algorithm.*

Key Words: **Machine Learning, Stocks, Stock Prediction, Deep Learning, Prophet, LSTM.**

1. INTRODUCTION

A stock market or share market is the accumulation of buyers and sellers with the stocks which represent ownership on businesses. The first and foremost aim of any investor before investing in any stock is to get significant profit. The age of economic growth and the endowment of digital technology led to the agglomeration of monetary data. The rapid growth of data ruled out the capability of humans to assess them manually. The inexhaustible expansion of such enormous oscillating and inconsistent data has put forward the demand to encourage an automated approach to financial data.

The stock market is one of the most important ways for any company to raise funds. Stock price forecasting always houses the risk but the profit has a wide effect in the field of economics and investments. Stock market prediction is the process of trying to determine the value of any particular stock in the future days. This is done by taking into consideration the previous stock prices and the variation of those prices in the previous days.

1.1 Machine Learning

Machine learning is defined as a technique of analyzing the data that automatizes the systematic model establishment. ML is a branch of artificial intelligence that works on the idea that systems can learn from data, patterns and make decisions. This method of learning requires minimum human intervention. Machine learning is described using 3 parameters that is P, E, T where T is

the task learned, E is the experience through which T is learned and performance P varies with E.
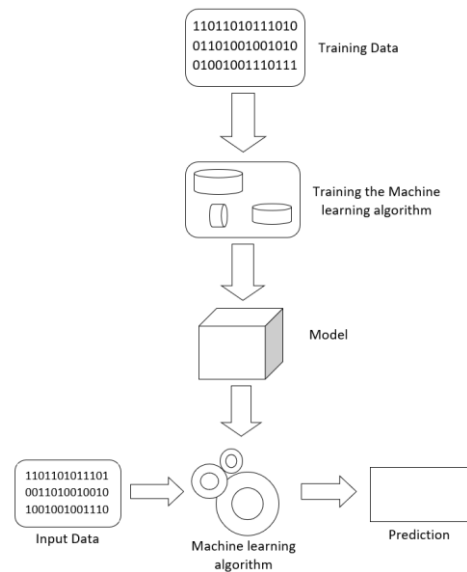

Fig – 1: Machine Learning

1.2 Deep Learning

Deep learning is a subset of machine learning in artificial intelligence (AI) that uses multiple layers to progressively extract higher-level features from raw input. In deep learning, each learning level transforms the input data given into more abstract and composite representation. Learning can be supervised, semi-supervised or unsupervised. Deep learning has been widely used in financial areas such as stock market prediction, financial information processing, and trade execution strategies.
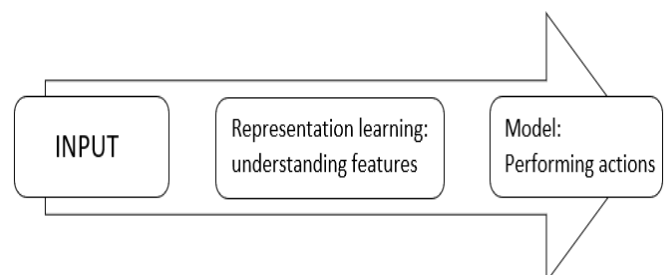

Fig – 2: Deep Learning

2. METHODOLOGIES

Moving Average

It uses the latest set of values for each prediction. For each subsequent step, the predicted values are taken into consideration while removing the oldest observation. The predicted closing price for a particular day will be the average of a set of previously observed values.
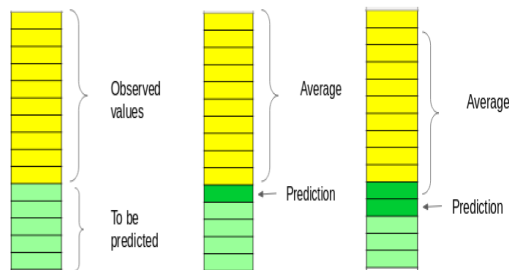

Fig – 3: Moving average

Linear Regression

It is the most basic machine learning algorithm that can be implemented. This algorithm returns an equation that determines the relationship between the independent variables and the dependent variable.

The equation for linear regression can be written as:
$$Y = \theta_1 X_1 + \theta_2 X_2 + \cdots \theta_n X_n$$
Here, x_1, x_2, ...x_n represent the independent variables while the coefficients θ_1, θ_2, ... θ_n represent the weights.

Support vector Machine

It is a supervised machine learning model that uses algorithms for classifying two-group classification problems. The main aim of the svm algorithm is to find a hyperplane in an N-dimensional space that distinctly classifies the data points.

Auto ARIMA

It is a very popular statistical method for time series forecasting. It stands for Autoregressive Integrated Moving average. ARIMA model takes into account the past values to predict the future values. It is a form of regression analysis that takes the strength of one dependent variable related to other changing variables.

Prophet

Prophet is a time series forecasting library that requires no data preprocessing and is extremely easy to implement. It is designed and pioneered by Facebook. Prophet is vigorous to missing data and shifts in trend. It

handles outliners well and also works best with time series that have strong seasonal effects.

Long Short Term Memory (LSTM)

LSTMs are a special kind of recurrent neural network, capable of learning long-term dependencies. It is used in the field of Deep Learning. LSTMs are widely used for sequence prediction problems and have proven to be effective. LSTM can store past important information, and forget the one that is not.

3. LITERATURE SURVEY

[1] Survey on stock market prediction using machine learning techniques

The objective is to predict the market performance with the help of an artificial neural network. The techniques of artificial neural networks classify the stock in mainly three categories that is buy, hold and sell, supported previous data. It's observed that the logistic regression model is employed by every individual to predict a stock in an exceedingly better way.

[2] Survey on stock market prediction using machine learning algorithms.

The objective is to predict the stock movement with good accuracy. The realm of interest is that the dataset of the stock prices from past years. The raw dataset must be pre-processed for data analysis. After pre-processing the information, we are going to use machine learning techniques like random forest and support vector machines on the dataset to come up with the outcomes.

[3] Survey on predicting stock prices using LSTM

Forecasting the stock prices has always been a difficult task for the analysts. The investors are highly curious about the stock prediction. For a successful investment, many investors have an interest in knowing the long-run condition of the stock market. The prediction models for the stock market helps the investors and also the analysts by providing the long run information of the stock market. Recurrent neural networks (RNN) and Long Short-Term Memory (LSTM) is the machine learning approaches used for stock price prediction.

[4] Survey on stock trend prediction using regression Analysis

A data mining approach various organizations are collecting data, building large data warehouses to store the collected data. Discovering the knowledge out of the collected data is done by employing a technique called data mining. Data mining software tool is employed to extract values of the

variables from the dataset to predict the long-run values of other variables with the utilization of time series data.

[5] Survey on stock market prediction using machine learning

Developing an application for analysing and predicting stock market prices increases the investor's interest in stock markets. First, we have to analyse the present and emerging methods of stock price prediction. The different approaches are fundamental analysis, technical analysis, and also the application of machine learning. Fundamental analysis and machine learning are accustomed to guide an investor's decisions. Whereas the technical analysis methodology provides limited useful information.

[6] Survey on stock market prediction: using historical data analysis

The stock market plays a vital role in business and finance. Fundamental analysis is depleted using social media data with the help of Sentimental Analysis. Social media data includes a high impact because of its increased usage, and it helps predict the trend of the stock market. Technical analysis is depleted by applying machine learning algorithms on historical data of stock prices. The tactic usually involves gathering various social media data, news to extract sentiments expressed by individuals.

4. OBSERVATIONS

SL.NO	IMPLEMENTATIONS	CONS	PROPOSED SOLUTION
1.	Artificial neural network with backpropagation algorithm	Neither growth nor pruning methods were attempted for the selection of network architecture.	Applying ANN tests on real-time data. Implementing ANN with inputs spread over the entire market instead of concentrating on a particular sector.
2.	Random forest Algorithms, support vector machine	Previous years dataset is considered. No real-time data are used for predicting stocks.	Considering more parameters to obtain higher accuracy. These Algorithms are implemented on public comments to understand the relationship between customer and employee.
3.	Root Mean Square Error (RMSE), the difference between the target value and the obtained output value is reduced by using RMSE value. Recurrent Neural Network, Long Short-Term Memory	Doesn't focus on events in the environment, like news or social media. It exploits only one data source, thus it is highly biased.	The future enhancement includes comparing the accuracy of LSTM with other prediction algorithms. LSTM gives more accurate value when compared with other prediction algorithms.
4.	Linear regression, moving average	Used for limited company stocks More amount of data is not considered for prediction	With the moving average algorithm, it is shown that the algorithm understands

			the past data and does not focus on the seasonal part. Therefore, accuracy is more.
5.	Artificial neural network, multiple linear regression, Bayesian Algorithm	using Bayes theorem bias is found. Predicted price is fluctuating they are not constant	As Bayes theorem provides bias in stock prices so the focus of the Algorithm can be moved to Ann using this seasonal stock prices can be predicted.
6.	SVM, ANN SVM (Support vector Machine)	Only sentiment data are used from various news and Twitter resources no historical data are considered for predictions.	The algorithm used for sentiment analysis uses a summative assessment of the sentiments in a particular tweet or news article, by categorizing them as positive, negative, and neutral data. The accuracy of the prediction could be improved by combining historical along with sentimental data.
7.	LSTM Neural Network Algorithm	Simply by considering the effect of historical data on price movements is too singular and may not be able to completely and precisely forecast the price on a particular day.	Adding data predictions related to stock-related news and basic information, to enhance the stability and accuracy of the model in the case of a major event.
8.	ARIMA, Facebook Prophet, Recurrent Neural Network LSTM	Models did not perform well in cases where stock prices are low or highly volatile. The models that used text (financial news articles) as part of the input have performed very well, while models that predicted future stock prices by historical stock prices lead to high percentage errors.	To verify if there is any effect or influence on the stock price of a particular company due to the stock price changes of other companies. By using both sentiments and historical data we can predict the accurate stock price of a company.
9.	Regression Based Model, LSTM	Sentiment analysis is not used. The larger dataset is not considered.	In the future, the accuracy of the stock market prediction system can be

			further improved by utilizing a much bigger dataset than the one being utilized currently. Sentiment analysis through Machine Learning on how news affects the stock prices of a company is also a very promising area.
10.	Random forest, SVM, sentiment analyzer tools	Considered only twitter data for sentimental analysis hence, not all the people who trade in stocks share their opinions on twitter.	Further other news resources for sentiment analysis can be used. Along with this historical data can be used to obtain high accuracy.

5. CONCLUSION

In this survey paper we conclude that though various methods and approaches can be used for predicting the stock price, every method has its limitations and advantages. By considering both technical indicators and fundamental analysis the accuracy of the predictions can be made more accurate and reliable. Despite many such algorithms available, there is always room for improvement. It is observed that use of different parameters of the data set in different algorithms results in various accuracy rates. We conclude that the use of time series forecasting with the prophet and artificial recurrent neural network that is LSTM (long short term memory) yield in more accurate prediction.

REFERENCES

[1] Saurav Agrawal, Dev Thakkar, Dhruvil Soni, Krunal Bhimani, Dr. Chirag Patel, "Stock Market Prediction using Machine Learning Techniques". International Journal of Scientific Research in Computer Science, Engineering and Information Technology, Volume 5, Issue 2, 2019. ISSN: 2456-3307, DOI: https://doi.org/10.32628/CSEITI952296.

[2] K. Hiba Sadia, Aditya Sharma, Adarsh Paul, SarmisthaPadhi, Saurav Sanyal, "Stock Market Prediction Using Machine Learning Algorithms". International Journal of Engineering and Advanced Technology, Volume 8, Issue 4, April-2019. ISSN: 2249-8958.

[3] Murtaza Roondiwala, Harshal Patel, Shraddha Varma, "Predicting Stock Prices Using LSTM". International Journal of Science and Research 2017. ISSN: 2319-7064

[4] S Abdulsalam Sulaiman Olaniyi, Adewole, Kayode S, Jimoh, R. G, "Stock Trend Prediction Using Regression Analysis – A Data Mining Approach". ARPN Journal of Systems and Software, Volume 1, Issue 4, 2011. ISSN: 2222-9833

[5] Gareja Pradip, Chitrak Bari, J. Shiva Nandhini, "Stock market prediction using machine learning". International Journal of Advance Research and Development, Volume 3, Issue 10, 2018.

[6] Vivek Kanade, Bhausaheb Devikar, Sayali Phadatare, Pranali Munde, "Stock market prediction: Using historic data analysis". International journal of advanced research in computer science and software engineering, volume 7, issue 1, 2017. ISSN: 2277 128X. DOI: 10.23956/ijarcsse/V711/0112.

[7] Dou Wei, "Prediction of stock price based on LSTM neural network". International conference on artificial intelligence and advanced manufacturing (AIAM) 2019. DOI: 10.1109/AIAM48774.2019.00113

[8] Saloni Mohan, Sahitya Mullapudi, Sudheer Sammeta, Parag Vijayvergia, David C Anastasiu, "Stock price prediction using news sentiment analysis". IEEE fifth international conference on big data computing service and applications 2019. DOI:10.1109/BidDataService.2019.00035

[9] Ishita Parmar, NAvanshu Agarwal, Sheirsh Saxena, Ridam Arora, Shikhin Gupta, Himanshu Dhiman, Lokesh Chouhan, "Stock market prediction using machine learning". International conference on secure cyber computing and communication (ICSCCC) 2018. 978-1-5386-6373-8/18/$31.00

[10] Venkata Sasank Pagolu, Kamal Nayan Reddy Challa, Ganapati Panda, Babita Majhi, "Sentiment analysis of twitter data for predicting stock market movements". International conference on signal processing, communication, power and embedded system (SCOPES) 2016. 978-1-5090-4620-1/16/$31.00