BANGALORE INSTITUTE OF TECHNOLOGY

K.R Road, V.V Pura, Bengaluru-04

DEPARTMENT OF ARTIFICIAL INTELLIGENCE & MACHINE LEARNING

PID-30

"Microsoft Stock Price Prediction"

Presented by

Meghana SR

1BI20AI031

VII Semester

Under the Guidance of

Prof. Subha Meenakshi S

Introduction

Stock: Stock refers to a share in the ownership of a company. Owning a stock represents a claim on part of the company's assets and earnings

Stock Market: The stock market refers to the collection of markets and exchanges where regular activities of buying, selling, and issuance of shares of publicly-held companies take place.

Such financial activities are conducted through institutionalized formal exchanges or over-the-counter (OTC) marketplaces which operate under a defined set of regulations.

Stock market analysis is divided into two parts :

- **Fundamental Analysis** involves analyzing the company's future profitability on the basis of its current business environment and financial performance.
- **Technical Analysis**, on the other hand, includes reading the charts and using statistical figures to identify the trends in the stock market.

Problem Statement

“To predict the stock price using dataset of Microsoft stock prices from April 2015 to April 2022.”

S.NO.	Authors name & Year of Publication	Title name and Journal name	Objective	Limitations
1.	Manas Kantimahanti, Bharat Thakkar, Kanishka Bisen, and Prof. Ramesh Mali,2022	Stock Market Prediction Using Neural Networks, International Research Journal of Engineering and Technology (IRJET)	Stock market prediction model using neural networks to analyze and predict stock prices.	Challenges in accurately capturing complex market dynamics
2	Tao Xing, Yuan Sun, Qian Wang, Guo Yu, 2013	Stock Price Prediction Using LSTM, Test Engineering and Management	Uses of multi-source, multiple-instance learning methods to forecast securities exchange.	Challenges in dealing with high-frequency trading, and unforeseen market events
3	Payal Soni, Yogya Tewari, and Prof. Deepa Krishnan,2020	Machine Learning Approaches in Stock Price Prediction, Journal of Physics: Conference Series	It aims to provide a systematic review of machine learning approaches in stock price prediction.	Limited training sets of financial information, which makes knowledge graph extraction challenging
4	Ishita Parmar,Ridam Arora Lokesh Chouhan,Navanshu Agarwal,Shikhin Gupta, Sheirsh Saxena,Himanshu Dhiman,2018	Stock Market Prediction Using Machine Learning, National Institute of Technology Hamirpur, India	Predict the future value of the financial stocks of a company using machine learning	Challenges in using larger dataset and provides lesser accuracy

Proposed System

The proposed system for predicting Microsoft stock prices adopts a comprehensive approach by incorporating diverse machine learning models to optimize forecasting accuracy

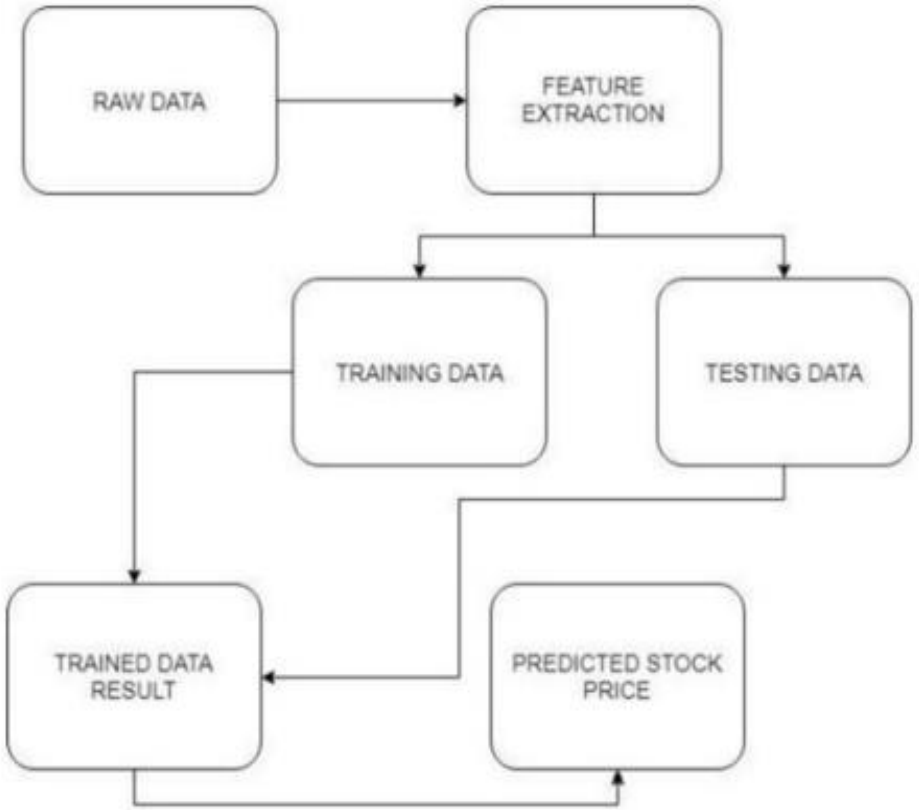
>>Linear Regression

>>Moving Average

>>Auto ARIMA

>> K-nearest neighbors

>>LSTM



Understanding the dataset:

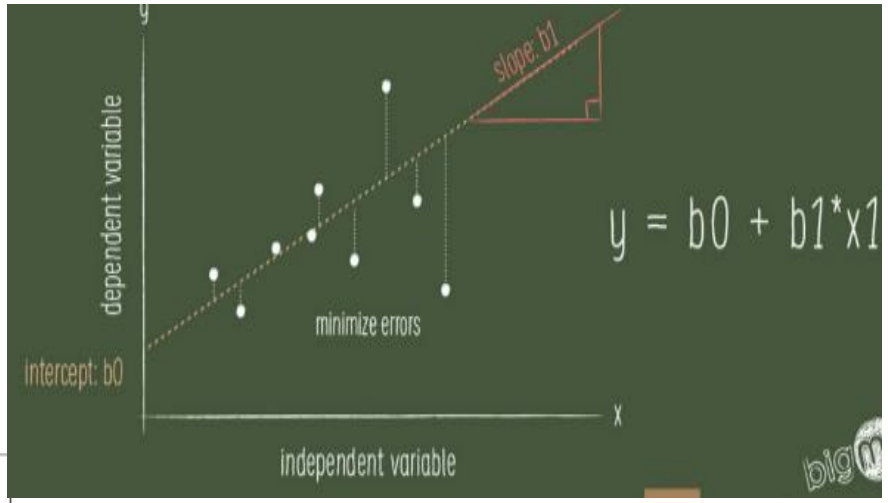
- There are multiple variables in the dataset – Date, Open, High, Low, Close and volume.
- The columns Open and Close represent the starting and final price at which the stock is traded on a particular day.
- High and Low represent the maximum and minimum price of the share for the day.
- Volume is the number of shares bought or sold in the day
- Market is closed on weekends and public holidays

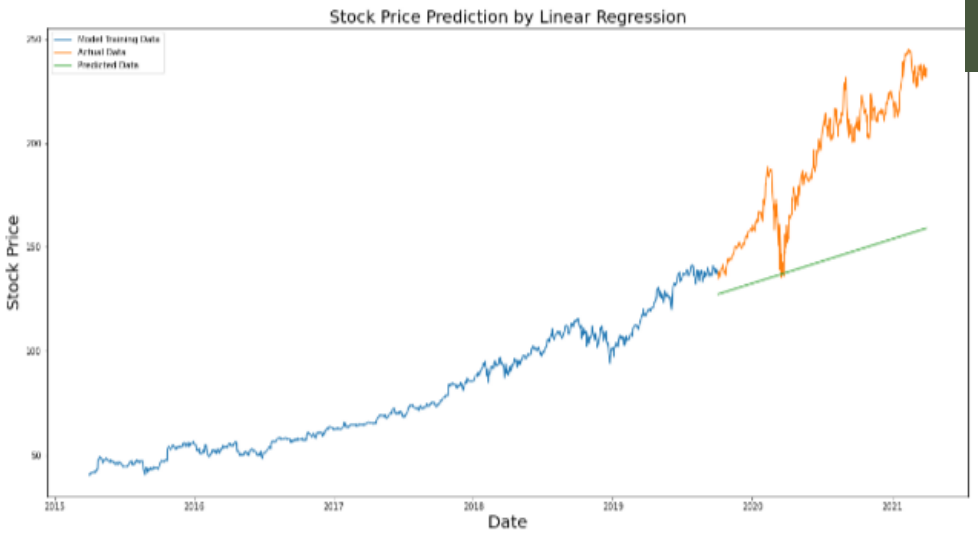
	Date	Open	High	Low	Close	Volume
0	4/1/2015 16:00:00	40.60	40.76	40.31	40.72	36865322
1	4/2/2015 16:00:00	40.66	40.74	40.12	40.29	37487476
2	4/6/2015 16:00:00	40.34	41.78	40.18	41.55	39223692
3	4/7/2015 16:00:00	41.61	41.91	41.31	41.53	28809375
4	4/8/2015 16:00:00	41.48	41.69	41.04	41.42	24753438
5	4/9/2015 16:00:00	41.25	41.62	41.25	41.48	25723861
6	4/10/2015 16:00:00	41.63	41.95	41.41	41.72	28022002
7	4/13/2015 16:00:00	41.40	42.06	41.39	41.76	30276692

Linear Regression

The most basic machine learning algorithm that can be implemented on this data is linear regression. The linear regression model returns an equation that determines the relationship between the independent variables and the dependent variable

Here, x1, x2,….xn represent the independent variables while the coefficients θ1, θ2, …. θn represent the weights.



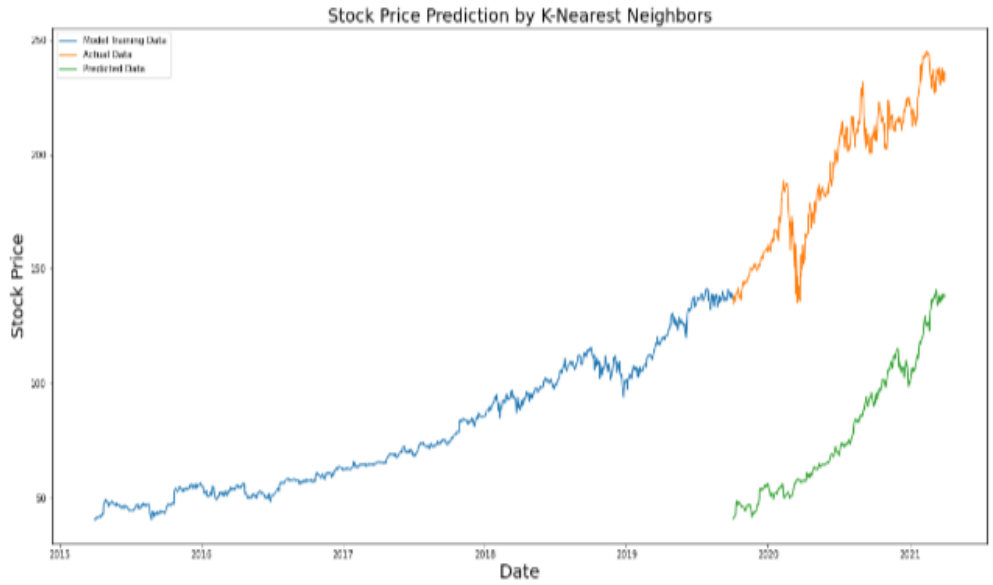


K-Nearest Neighbours

Based on the independent variables, KNN finds the similarity between new data points and old data points.

- The RMSE value is almost similar to the linear regression model and the plot shows the same pattern. Like linear regression, KNN also identified a drop in January 2018 since that has been the pattern for the past years.
- Regression algorithms have not performed well on this dataset.

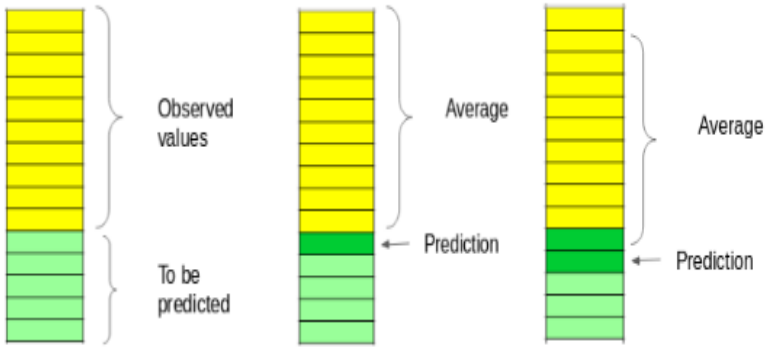
RMSE: one of the two main performance indicators for a regression model. It measures the average difference between values predicted by a model and the actual values.

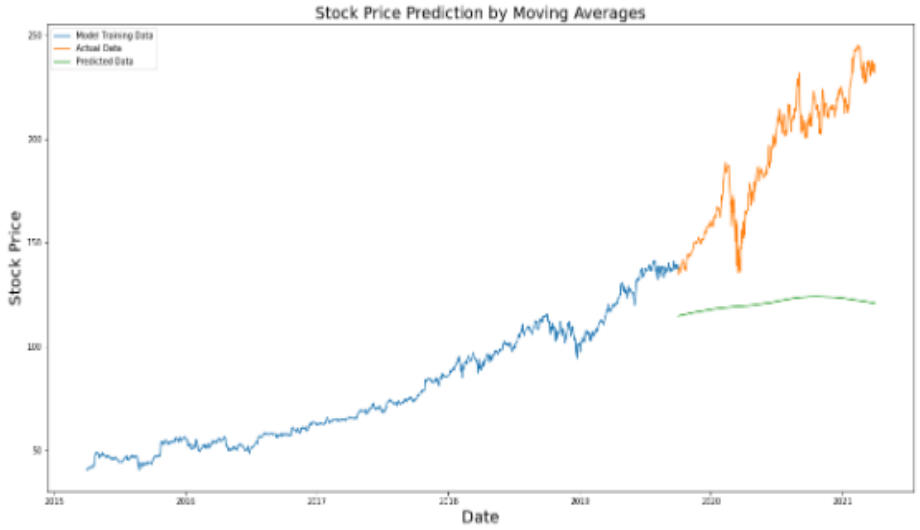


Moving Average

‘Average’ is easily one of the most common things we use in our lives.(eg: calculating marks, temp, performance)

The predicted closing price for each day will be the average of a set of previously observed values. Instead of using the simple average, we will be using the moving average technique which uses the latest set of values for each prediction. In other words, for each subsequent step, the predicted values are taken into consideration while removing the oldest observed value from the set.

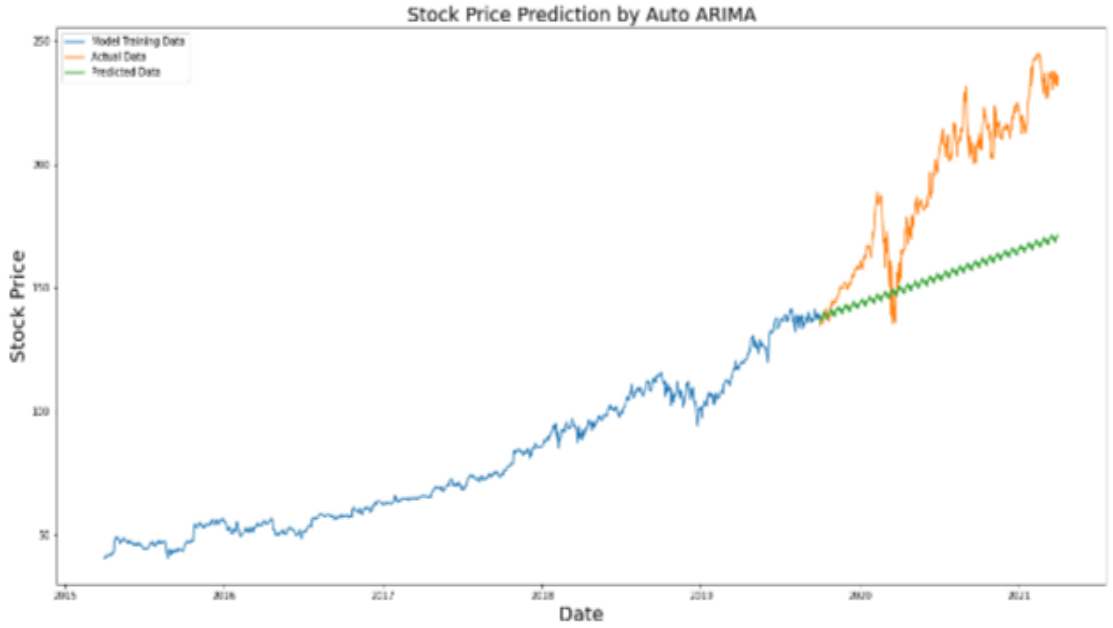




Auto ARIMA

ARIMA is a very popular statistical method for time series forecasting. ARIMA models take into account the past values to predict the future values. There are three important parameters in ARIMA:

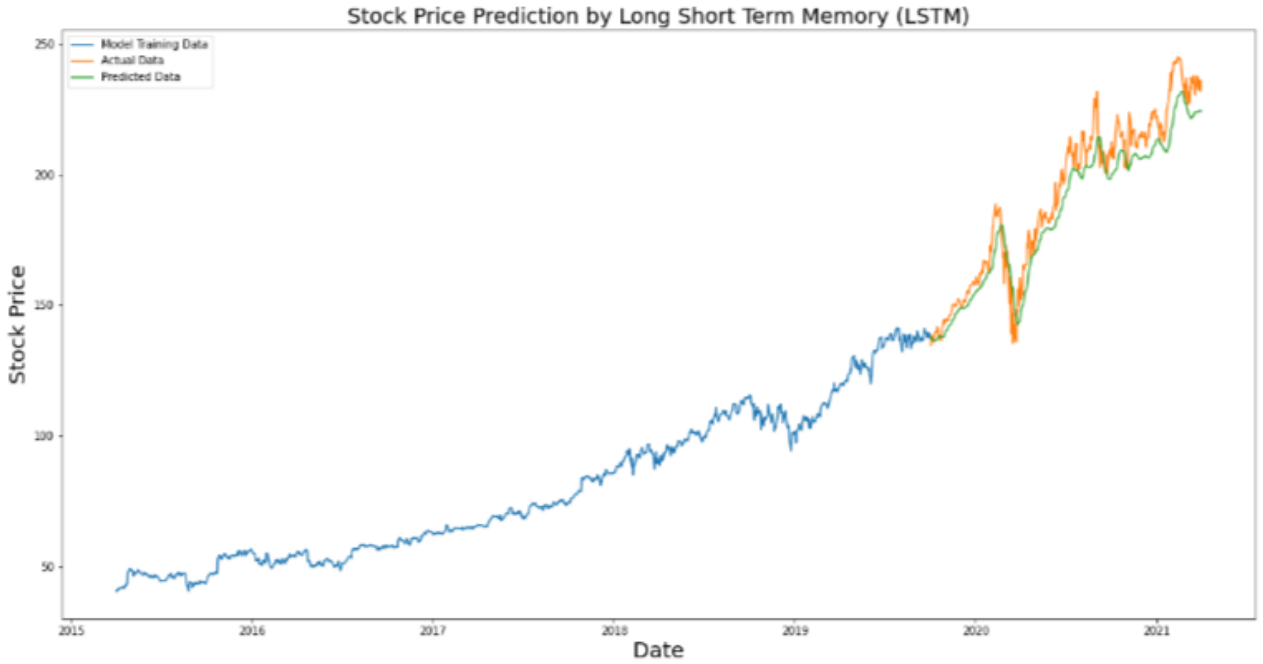
- p (past values used for forecasting the next value)
- q (past forecast errors used to predict the future values)
- d (order of differencing)



Long Short Term Memory (LSTM)

LSTMs are widely used for sequence prediction problems and have proven to be extremely effective. The reason they work so well is because LSTM is able to store past information that is important, and forget the information that is not. LSTM has three gates

- The input gate: The input gate adds information to the cell state
- The forget gate: It removes the information that is no longer required by the model
- The output gate: Output Gate at LSTM selects the information to be shown as output



CONCLUSION

- In predicting Microsoft's stock price, four distinct methodologies were employed: linear regression, moving average,KNN and LSTM (Long Short-Term Memory).
- It's essential to note that the efficacy of these methods is contingent on the dynamic and multifaceted nature of financial markets, and predictions should be interpreted with caution due to the inherent uncertainty in stock price movements.

Thank you!!