STOCK PRICE PREDICTION

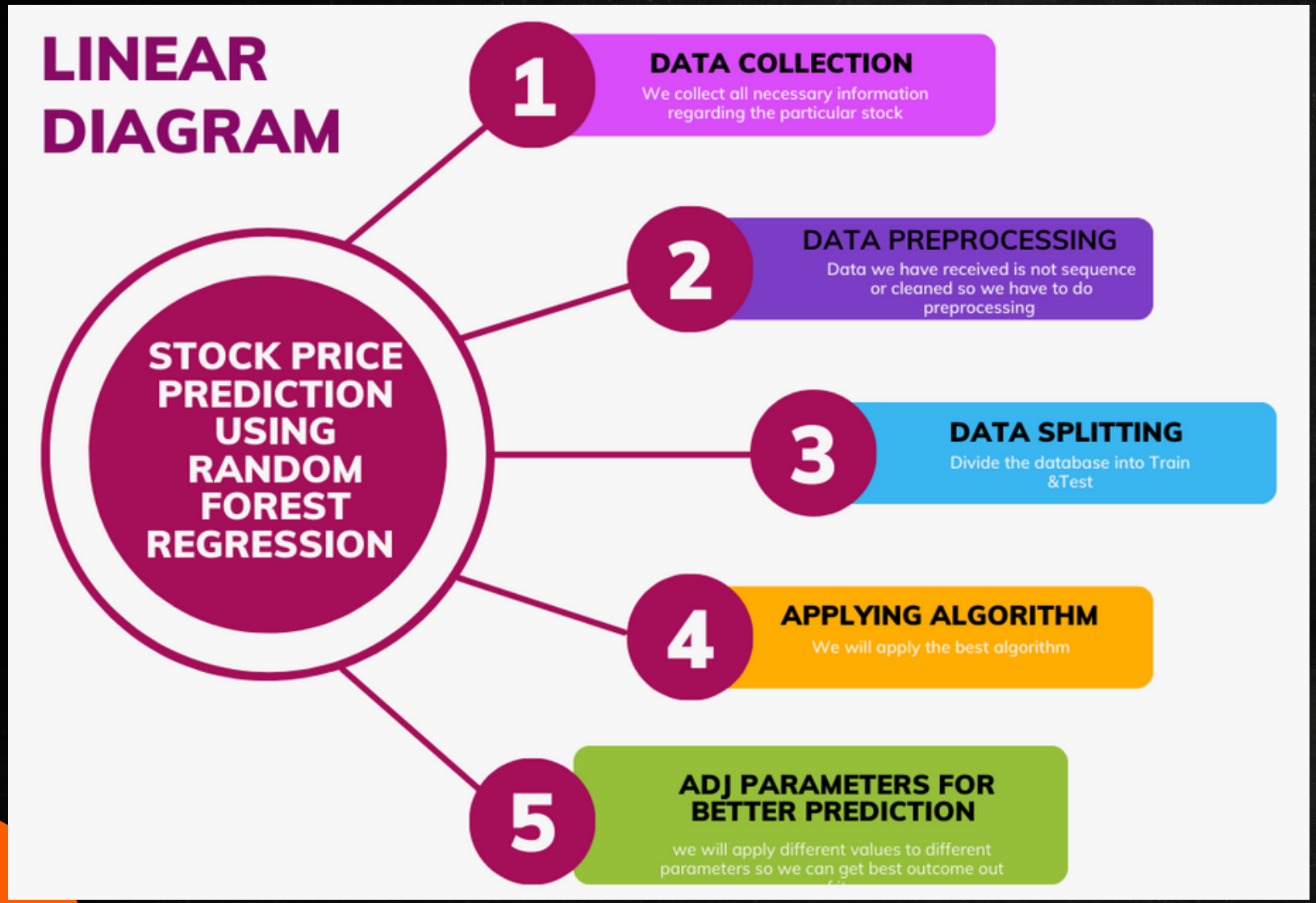
GUIDED BY : MR.DEEPAK GURAV SIR
PRESENT BY : ATHARV DARP

CONCEPT

In this study, i delve into the application of the Random Forest algorithm for the purpose of stock price prediction. The stock market is known for its inherent volatility and complexity, making accurate predictions a challenging task. We take historical stock data, encompassing information such as past prices, trading volumes, and relevant financial indicators, to construct predictive models.

Random Forest, chosen for its robustness and ability to handle intricate datasets, emerges as a promising candidate in this context. It operates by aggregating predictions from a multitude of decision trees, reduce the risk of overfitting and accommodating non-linear relationships within the data. This approach holds the potential to enhance forecasting accuracy, a paramount concern for both seasoned investors and newcomers to the stock market.

STRUCTURE OF MODEL



LOOPHOLES

Market Volatility: May struggle to predict extreme market fluctuations during periods of high volatility.

Non-Linear Patterns: Difficulty capturing highly non-linear stock price movements.

Data Quality: Predictions are sensitive to the accuracy and quality of historical data.

Changing Market Conditions: Limited adaptability to new market conditions and unforeseen events.

Overfitting: Risk of overfitting to historical data, leading to poor generalization.

Feature Selection: The choice of features can significantly affect model performanc

PROBLEMS IDENTIFICATION

"The main problem we aim to address is the difficulty in predicting stock prices accurately, which arises due to factors such as market volatility, non-linearity in price movements, the influence of external events, and the challenge of incorporating these complex dynamics into reliable predictive models."

addition to the complexities mentioned, we also face the issue of data quality and availability. Stock price prediction relies heavily on historical data, and inconsistencies or missing information in these datasets can lead to unreliable forecasts

OUR OBJECTIVE BEHIND CREATING THIS MODEL

The primary objective of stock price prediction is to provide investors and financial professionals with actionable insights into the future movements of stock prices. By harnessing historical price data, market indicators, and advanced predictive models

Our goal is to reduce uncertainty and enhance decision-making, empowering individuals and institutions to optimize investment portfolios, minimize risks, and capitalize on market opportunities. Stock price prediction represents the convergence of data science, finance, and technology, aiming to make the unpredictable nature of financial markets more manageable.

CONCLUSION

At conclusion, using a Random Forest regression model for stock price prediction is a valuable tool, but it has limitations related to market volatility, data quality, and model complexity. To address these limitations, it's essential to combine it with other methods, carefully select features, and incorporate risk management strategies in your investment approach.

ABOUT RANDOM FOREST REGRESSION

Random Forest Regression: Random Forest is a powerful machine learning algorithm that combines multiple decision trees to make predictions. In regression tasks like stock price prediction, it aggregates predictions from many trees to provide more accurate and robust forecasts. It's known for handling complex relationships in data and mitigating overfitting.

Also I have applied another regression model also but even after using hyper parameter tuning model get overfited so decided to use this model

RANDOM FOREST REGRESSION PARAMETERS

n_estimators :- This parameter defines the number of decision trees in the Random Forest. A higher number of trees can improve the model's accuracy, but it may also increase computation time. Common values are in the range of 100 to 1000.

max_depth: It determines the maximum depth or level of each decision tree. Deeper trees can capture more complex patterns in the data but are more prone to overfitting. You can set this parameter to limit tree depth and control overfitting.

K - Fold :- when the dataset is split into a K number of folds and is used to evaluate the model's ability when given new data.

RMSE :- it is square-root of MSE
MSE :- MSE is RMSE is difference between actual value and predicted value
Bagging and Boosting - Bagging help to

reduce variances to week learners , while Boosting use reduce bios for week learners

OUR DATASET

DATE :- DATE Column's describe about the on which date price moved (Actual Traded Day)

Open :- Open Column's describe about Open price of that candle

HIGH :- HIGH Column's describe about that highest price of that stock

LOW :- LOW Column's describe about that candle's lowest price

VOLUME :- VOLUME Column tell about in how much quantity that stock has traded (Liquidity of that stock)

THANK YOU!!