**TITLE: Stock Price Prediction**

**Abstract:**

In Stock price Prediction, aim is to predict the future stocks prices of a company. The recent trend in stock price prediction technologies is the use of machine learning which makes predictions based on the values of current stock market indices by training on their previous values. Machine learning itself employs different models to make prediction easier and authentic. we focused on the use of Regression (random forest algorithm) based Machine learning to predict stock values. Factors considered are open, close, low, high and volume.

**Problem Statement:**

The problem is to build a predictive model that forecasts stock prices based on historical market data. The goal is to create a tool that assists investors in making well-informed decisions and optimizing their investment strategies. This project involves data collection, data preprocessing, feature engineering, model selection, training, and evaluation.

**Design Thinking:**

1, Understanding of the needs and challenges of traders , investors ,and analysts who rely on stock price predictions.

2, Define the problem statement and goals of the stock price prediction model. Identify specific aspects that need improvement, such as accuracy, timeliness, or interpretability.

3, Generate ideas for potential solutions. Consider various approaches, such as machine learning algorithms, technical analysis techniques.

4, Create a prototype of the stock price prediction model using suitable tools and technologies. Test different algorithms, data sources, and indicators to refine the model.

5, Evaluate the prototype using historical data and performance metrics.

6, Develop a production-ready stock price prediction model based on the refined prototype. Ensure scalability, reliability, and real-time functionality.

7, Deploy the model in a suitable environment, such as a web-based platform or a trading software.

8, Continuously monitor and update the stock price prediction model based on real-time data and user feedback. Incorporate new techniques and improve upon existing algorithms to enhance accuracy and reliability.

**Phase of Development**:

1,Data Collection:

In these project , first step involves collecting historical

Financial market data , including features such as stock’s open high low close prices as well as volume….The more comprehensive and accurate the data, the better the predictions can be . It’s important to ensure the data is clean before using it for training the ml models.

2,Data Preprocessing:

Once the data is collected,it needs to be cleaned and preprocessed. Clean the data by handling missing values, removing outliers and normalizing the features (transforming the data into a suitable format for ml algorithms).Feature engineering techniques can also be applied to extract features from raw data.

3,Model Selection:

In these selecting a model, I selected a random forest algorithms,

Random forest : Random forest is an ensemble learning algorithm that combines multiple decision trees to make predictions.

It can be used in stock price prediction by utilize historical stock data as input and future stock price as the target variable

4, Model Training:

Split the data into training and testing sets. Train the model using the training data.

5, Model Evaluation:

Once trained the model is evaluated on the testing data to assess its performance

Evaluation metrics for stock price prediction include mean squared error (MSE) root mean squared error (RMSE) mean absolute error (MAE)..

6, Prediction and Validation:

After evaluating the model , it can be used to make predictions on unseen data. The predicted stock prices compared to actual values for validation.

7, Model Deployment:

Deploy the model to a web application in a real-time

8, Monitoring and Maintainance:

Continuously monitor the model’s performance and retrain it periodically with fresh data.Update the model with new data to keep predictions accurate.

**Dataset:**

Features like date,adj close,close,open,high,low,volume……

Date: Date of trade

Open: Opening price of the stock

High: Highest price of the stock on that day

Low: Lowest price of the stock on that day

Close: Close price adjusted for splits

Adj Close: Adjusted close price adjusted for splits, dividends, and capital gain distributions

Volume: Volume of stock traded on that day

**Code:**

import pandas as pd

from sklearn.model\_selection import train\_test\_split

from sklearn.ensemble import RandomForestRegressor

from sklearn.metrics import r2\_score,mean\_squared\_error,mean\_absolute\_error

data = pd.read\_excel(r" C:\Users\admin\Downloads\ MSFT.xlsx")

date = data['Date']

date = pd.to\_datetime(date, format="%d-%m-%Y")

data['Year'] = date.dt.year

data['Month'] = date.dt.month

data['Day'] = date.dt.day

X = data[['Day','Month','Year','Open']]

y = data['Adj Close']

X\_train,X\_test,y\_train,y\_test = train\_test\_split(X,y,train\_size=0.2,random\_state=42)

model = RandomForestRegressor(n\_estimators=10,random\_state=42)

model.fit(X\_train,y\_train)

score = model.score(X\_test,y\_test)

prediction = model.predict(X\_test)

predicted\_score = r2\_score(y\_test,prediction)

mse = mean\_squared\_error(y\_test,prediction)

mae = mean\_absolute\_error(y\_test,prediction)

rmse = mean\_squared\_error(y\_test,prediction,squared=False)

print(f'Score : {score}')

print(f'R squared score : {predicted\_score}')

print(f'Mean squared error : {mse}')

print(f'Mean absolute error :{mae}')

print(f'Root mean squared error :{rmse}')

**Output:**

Score : 0.9994888994161154

R squared score : 0.9994888994161154

Mean squared error : 0.4098517797615241

Mean absolute error :0.31215766045454557

Root mean squared error :0.6401966727198167

**Conclusion:**

Thus we conclude that we predicted the stock price prediction by using random forest algorithm ,and also trained and evaluated..