ENHANCING INVESTMENT DECISIONS_PHASE1

PROBLEM TITLE : STOCK PRICE PREDICTION

INTRODUCTION

In the realm of finance, accurate stock price prediction is a sought-after advantage. Our project harnesses data science and machine learning to forecast stock prices, aiming to empower investors and traders with valuable insights in a complex and ever-changing market.

PROBLEM DEFINITION :

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.Data collection

https://www.kaggle.com/datasets/prasoonkottarathil/microsoft-lifetime-stocks-dataset

Using the above link we collect historical stock market date, open price, close price and other relevant indicators that is needed for our project to predict stock price.

2. Data preprocessing

Data preprocessing plays a critical role in stock price prediction models. We do steps like data cleaning and preprocessing the data which involves:

•Removing duplicate data entries if any exist.

•Identifying and handling missing data points by using techniques like interpolation or forward/backward filling.

•Checking for any outliers that may skew the data and decide on an appropriate treatment.

•Converting categorical features into numerical representations

•Normalization/Scaling, feature selection, handling imbalanced data and saving the preprocessed data are some of the important steps that we perform.

3.FEATURE ENGINEERING

Feature engineering is of paramount importance in stock price prediction for several reasons:

•Relevant features provide valuable information about the stock and its market environment

•Helps in model performance, reducing noise, risk management, overfitting prevention,etc..

Some steps and examples involved in feature engineering are:

•Historical Data: Use daily closing prices, high and low prices, and trading volume.

•Time-Based Features: Extract date-related information and calculate moving averages.

•Volatility Indicators: Include standard deviation and Bollinger Bands.

•Technical Indicators: Use RSI, MACD, and Stochastic Oscillator.

•Fundamental Data: Consider EPS, P/E ratio, and dividend yield if available.

•Market Sentiment: Analyze news sentiment and economic indicators.

•Lagged Features: Include lagged prices for temporal dependencies.

•Market Index Data: Incorporate market indices and their correlations.

•Seasonality: Capture recurring patterns.

4.MODEL SELECTION

Selecting the right model for stock price prediction is crucial for accurate and reliable forecasts. Here are some commonly used models and approaches for stock price prediction:

•ARIMA

•LSTM Neural Networks

•GRU Neural Networks

•Prophet, etc..

We can use Ensemble methods which combine multiple models, such as ARIMA, LSTM, or linear regression, into an ensemble to improve predictive accuracy.Techniques like stacking, bagging, or boosting can be applied to create ensemble models.

5.MODEL TRAINING

•We use preprocessed data to train the selected model (e.g., ARIMA, LSTM).

•Split data into training, validation, and test sets.

•Configure model architecture and hyperparameters.

•Train iteratively, monitoring key metrics.

•Implement early stopping and record training history for evaluation.

By doing these, the chosen model learns from historical data to make predictions about future stock prices.

6.EVALUATION

•We assess the performance of the trained stock price prediction model using appropriate time series forecasting metrics.

•Compute metrics such as Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) to quantify the model's accuracy in predicting stock prices.

•Utilize additional metrics like Mean Absolute Percentage Error (MAPE) or directional accuracy to gauge the model's reliability.

•Visualize the model's predictions against the actual stock price data to gain insights into its behavior.

Analyze any residual plots or error distributions to identify potential patterns or biases.

•Conduct statistical tests, if necessary, to validate the significance of the model's predictions.

•Consider comparing the model's performance against baseline models or alternative algorithms to assess its added value.

•Use the validation and test sets to ensure that the model generalizes well to unseen data.

•Document the evaluation results comprehensively, including any observations or insights gained during the analysis.

CONCLUSION:

Our project was designed to support investors and traders by providing valuable insights into stock price movements.This project underscores the capability of data science and machine learning in addressing intricate financial forecasting tasks. While absolute certainty in stock price prediction remains elusive, our work has yielded a resilient predictive tool, offering valuable guidance to market participants. The outcomes and methodologies establish a strong basis for advancing stock price prediction and algorithmic trading strategies.