

Project 1

Stock Price Prediction

Introduction:

This project aims to predict stock prices using machine learning techniques. The dataset contains historical stock market data, including daily stock prices, trading volume, and technical indicators.

Dataset

The dataset used for this project can be found on Kaggle: [Stock Market Prediction Dataset](#)

Project Overview

1. Exploratory Data Analysis (EDA):

Analyzed the structure and characteristics of the stock market dataset.

Visualized key statistics and trends in stock prices

2. Feature Engineering:

- Created moving average features (MA10 and MA50) to capture short-term and long-term trends.
- Generated lag features to include the previous day's closing prices.
- Calculated volatility and Relative Strength Index (RSI) as additional indicators.

3. Predictive Modeling:

- Implemented Linear Regression model.
- Utilized features generated from EDA and feature engineering steps for model training.
- Evaluated models using Mean Squared Error (MSE).

4. Documentation:

- Documented the approach, methodologies, and insights gained from the stock market dataset.
- Provided clear explanations for the chosen predictive models and feature engineering techniques.

Methodology

Exploratory Data Analysis (EDA)

- Analyzed the distribution of stock prices, trading volume, and technical indicators.
- Visualized trends and patterns in stock prices over time.
- Identified potential correlations between features and target variables (closing prices).

Feature Engineering

- Generated moving average features (MA10 and MA50) to capture short-term and long-term trends in stock prices.
- Created lag features to include historical closing prices.
- Calculated volatility and RSI as additional indicators to capture market momentum.

Predictive Modeling

- Utilized Linear Regression model for stock price prediction.
- Split the dataset into training and testing sets.
- Trained the models using the training data and evaluated their performance using Mean Squared Error (MSE).

Results

- **Linear Regression:**

Achieved a Mean Squared Error (MSE) of **13.9867877** on the testing set.

Conclusion

- The project successfully implemented machine learning techniques to predict stock prices based on historical data.
- Feature engineering techniques such as generating lag features, moving averages, and additional indicators like volatility and RSI improved model performance.

Future Work

- Explore more advanced machine learning algorithms such as Gradient Boosting Machines (GBM) or Long Short-Term Memory (LSTM) networks for further improvement.
- Incorporate additional external data sources such as news sentiment analysis or macroeconomic indicators for more robust predictions.