# Related Work on Stock Market Prediction

Stock market prediction has been a popular topic in the field of finance, data science, and machine learning. Researchers have used various techniques and methods to predict stock prices and trends. Below is an overview of related work and how the current project is different:

**1. Linear Regression for Stock Prediction**

- Previous projects often use simple linear regression to model the relationship between stock features such as Open, High, Low, and Close prices. Linear regression is easy to understand but cannot capture the complex patterns in stock data.

- Difference: This project uses additional models like Support Vector Regression (SVR) and Random Forest to handle non-linear relationships in the data. Combining regression and classification models helps create a more complete prediction approach.

**2. Time-Series Forecasting with ARIMA**

- Many studies use ***ARIMA (Auto-Regressive Integrated Moving Average)*** to forecast stock prices based on past data. ARIMA models focus on time trends but may miss other important factors like volume and seasonality.

- Difference: In this project, we include features like Volume, Month, Day, and Year to capture more factors that can influence stock price movements. This helps to model the data more effectively.

**3. Machine Learning Classifiers for Price Movement**

- Some projects use classification algorithms like Logistic Regression or Random Forest to predict if the stock price will go up or down. These models give a yes/no answer based on previous closing values.

- Difference: In this project, we use a mix of classifiers, including Naive Bayes, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN). We also use different metrics and visualizations, such as confusion matrices, to evaluate the performance of each model.

**4. Neural Networks for Stock Prediction**

- Recently, deep learning models like Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks have become popular for stock prediction. These models are powerful for handling sequences of data and perform well with large datasets.

- Difference: Instead of deep learning, this project uses traditional machine learning models to make it easier to understand and implement. The focus is on using simpler models that are computationally less expensive and more accessible.

**5. Sentiment Analysis-Based Predictions**

- Some projects also use sentiment analysis on news articles or social media to understand market sentiment and predict stock price movements. These methods use natural language processing to add extra information.

- Difference: This project focuses only on structured financial data, such as historical prices and volume. This makes the project easier to reproduce and avoids the complexity of dealing with text data.

## Summary of Differences

- **Model Diversity:** The current project uses a variety of models, including Linear Regression, SVR, KNN, Random Forest, and Polynomial Regression, to find the best-performing model.

- **Feature Engineering:** Unlike projects that focus solely on time-series prediction, we also use features derived from the date and volume, adding more information to improve the predictions.

- **Evaluation Techniques**: Each model is evaluated using scatter plots, regression lines, residual plots, and confusion matrices to better understand how well each model works.

- **Simplicity and Interpretability:** This project is designed to be understandable, using traditional machine learning techniques instead of complex deep learning models.

These differences show that this project aims to balance simplicity, interpretability, and performance by using a range of machine learning techniques to predict stock prices more effectively.