# Machine Learning Minor Project

By: Shelton Victor Coutinho

Batch: May 2025

**Dataset Description**

The dataset consists of historical stock price data, including features like **closing price** (how much the stock was worth at the end of each trading day) and **trading volume** (the number of shares traded). It was sourced from Yahoo Finance and covers multiple years.

**Algorithm Details**

* **K-Nearest Neighbors (KNN):** Compares stock prices with similar past patterns to make predictions. Works well for short-term forecasts.
* **Simple Linear Regression:** Tries to find a straight-line relationship between stock price and time or other features like trading volume. Good for basic trend analysis.
* **K-Means Clustering:** Groups stocks with similar behavior into clusters, helping investors understand price movement patterns.

**Accuracy**

* KNN typically performs well when there are clear historical patterns, but it may struggle with sudden price jumps.
  + Accuracy based on Output:
  + KNN Mean Absolute Error: 0.12076328444718198
* Linear Regression is simple but might not handle complex stock movements effectively.
  + Accuracy based on Output:
  + Linear Regression Mean Absolute Error: 0.12283739316563816
* K-Means clustering doesn’t directly predict prices but helps with grouping stocks based on similarities.
* Accuracy depends on market conditions, feature selection, and data quality. More advanced models like LSTMs can improve predictive power.