# Stock Market Prediction Using Machine Learning

## 1. Problem Statement

The objective of this project is to predict whether Google (GOOGL) stock price will increase the following trading day. This is framed as a binary classification problem where the target variable equals 1 if the next day's closing price is higher than today's, and 0 otherwise.

## 2. Data sources

Two external APIs were utilized to construct the dataset: Yahoo Finance, FRED API

Time Period: November 22, 2023 – November 22, 2025 (2 years)

Final dataset: 497 samples with 6 engineered features

## 3. Model: Decision Tree, Random Forest, XGBoost

## 4. Results

| Model | Accuracy | F1-Score | Precision |
|---|---|---|---|
| Decision Tree | 0.53 | 0.58 | 0.63 |
| Random Forest | 0.53 | 0.62 | 0.60 |
| XGBoost | 0.59 | 0.68 | 0.64 |

Best Model: XGBoost - 59% accuracy

## 5. Limitations

Efficient Market Hypothesis: Stock prices inherently incorporate all available information, making consistent prediction challenging.

Sample Size: 497 samples may be insufficient for complex models to learn robust patterns.

## 6. Next Steps

Expand Data Sources: Integrate news APIs (e.g. NewsAPI, GDELT), social media sentiment (Twitter/Reddit), and earnings announcement data.

Additional Technical Features: Implement RSI (Relative Strength Index), MACD (Moving Average Convergence Divergence), and other established technical indicators.