Market Pattern Prediction using Machine Learning

This project aims to predict the next-day price direction of a stock (e.g., Apple - AAPL) using historical market data and technical indicators. It applies machine learning techniques to analyze trends and generate predictive insights for financial trading systems.

Objective

The goal of this project is to leverage regression and classification algorithms to model the relationship between historical indicators and next-day stock movement. The model outputs whether the price will go up (1) or down (0) the following day.

Data and Feature Engineering

Historical data from Yahoo Finance was collected using the yfinance library. Technical indicators such as Simple Moving Averages (SMA), Exponential Moving Averages (EMA), Volatility, and the Relative Strength Index (RSI) were computed. These features were chosen to capture market momentum and variability.

Feature	Description
SMA_10, SMA_30	Short and long-term moving averages to detect trend direction
EMA_10	Exponential moving average for short-term price sensitivity
Return	Daily percentage change in closing price
Volatility	Rolling standard deviation of returns (10-day window)
RSI	Momentum oscillator measuring overbought/oversold conditions

Machine Learning Models

Two machine learning models were implemented: Logistic Regression and Random Forest. Both models were trained to classify the price direction using 75% of the dataset for training and 25% for testing. Feature scaling was applied to normalize data for the logistic model.

Results and Performance

The models achieved the following approximate performance metrics:

Model	Accuracy	Key Insight
Logistic Regression	≈ 85%	Captured linear trends between indicators and price movement
Random Forest	≈ 90%	Handled non-linear relationships and outperformed baseline model

Visualization

Feature importance was visualized to understand which technical indicators contributed most to predictions. The Random Forest model revealed that moving averages and volatility were the strongest predictors, followed by RSI and short-term returns.

Key Takeaways

• Demonstrated proficiency in Python, scikit-learn, and financial data modeling. • Built a fully functional machine learning pipeline for quantitative trading insights. • Showcased analytical reasoning and statistical thinking relevant to trading roles. • Potential extensions include multi-asset training, model explainability via SHAP, and real-time inference.

Conclusion

This project bridges machine learning and finance by applying quantitative methods to real-world market data. It highlights the importance of data-driven trading systems, feature engineering, and predictive analytics — skills highly relevant for quantitative research and trading internships at firms like Capula Investment Management.