

Stock Price Movement Prediction Using Random Forest

Samsung Innovation Campus — AI Project (Group 33)

Tharish J Naman Chawla Dina Abdelnaser Mohammed
Raia Mohammed Gyeye Nicelin Printo Mariam Lulu Mohammed

2025

Problem Statement

- Short-term stock price movement is **highly noisy** and influenced by:
 - news, macro events, sentiment, and investor behaviour.
- Traders often rely on a few manual indicators that can **fail in volatile markets**.
- There is no simple rule that reliably tells you if **tomorrow's close will be up or down**.

Core Question

Can we use data and AI to predict whether **tomorrow's closing price** will be UP (1) or DOWN (0) for a given stock?

Proposed Solution & AI Technique

Our idea: Build a machine-learning pipeline that learns patterns from historical data.

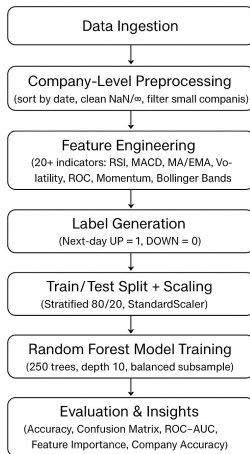
- Use daily OHLCV data (Open, High, Low, Close, Volume) for global stocks.
- Engineer **20+ technical indicators**:
 - RSI, MACD, EMA, Bollinger Bands
 - Momentum, Rate of Change (ROC)
 - Volatility, price range, daily returns, volume trends
- Predict next-day direction as a **binary classification** problem: UP / DOWN.

AI Model Used: Random Forest Classifier

- Ensemble of decision trees — robust to noise and non-linearity.
- Works well on tabular financial data.
- Provides **feature importance** for transparency.

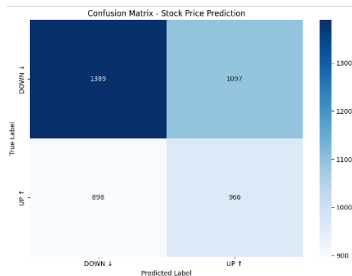
System Architecture & Workflow

System Architecture and Overview

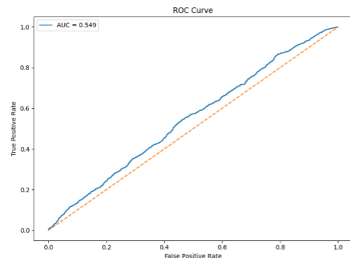


Results: Model Performance

- **Test accuracy:** 54% (baseline random = 50%).
- **ROC–AUC:** 0.549 $\Rightarrow$ model captures **weak but non-random** signals.
- Class-wise performance:
 - **DOWN days:** Precision 0.61, Recall 0.56, F1 = 0.58
 - **UP days:** Precision 0.47, Recall 0.52, F1 = 0.49



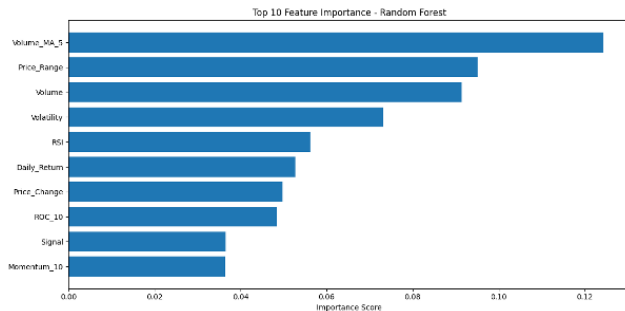
Confusion matrix



ROC curve (AUC = 0.549)

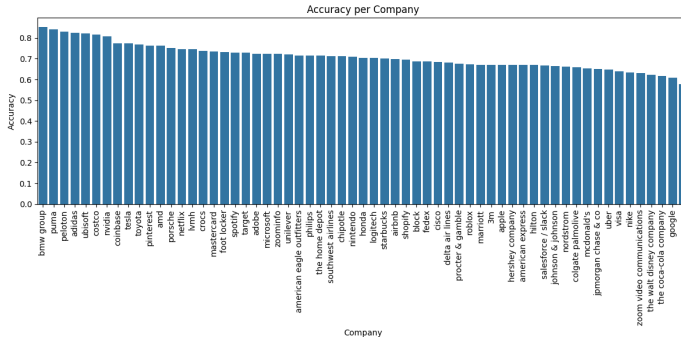
1. Which indicators matter most?

- Feature importance shows:
 - **Daily_Return, Volume_MA_5, Price_Change, Volatility**
 - are among the strongest predictors.



- Short-term direction is driven by:
 - price swings (range & returns),
 - volatility regimes,
 - short-term volume trends.
- Confidence scores cluster around **45–52%**:
 - signals are weak and noisy,
 - useful more for **analysis** than trading.

Key Insights: Company-Level Behaviour



- Some stocks are **much easier** to predict:
 - BMW, Puma, Peloton, Adidas, Ubisoft — accuracy ≥ 0.82 .
- Others behave almost randomly:
 - Google, Amazon, Coca-Cola — near 50% accuracy.
- Predictability is **stock-dependent**, not uniform across the market.

- Built an end-to-end pipeline for **next-day stock direction prediction**.
- Random Forest with engineered indicators reached **54% test accuracy**:
 - confirms that technical indicators contain **weak but usable** signals.
- Analysis showed:
 - volatility, returns, and volume trends are key drivers,
 - big variation in predictability across individual companies.
- The system is best viewed as a **research prototype**:
 - useful for learning and analysis,
 - not yet reliable for real-money trading.

Future Improvements

- **Richer data**

- add news & social media sentiment,
- include macroeconomic indicators and longer history.

- **Stronger models**

- try XGBoost / LightGBM,
- experiment with LSTM/GRU for sequential patterns.

- **Deployment**

- turn into a web dashboard,
- daily data updates and automatic retraining,
- allow users to select stocks and view predictions + explanations.

Thank you!

Questions and feedback welcome.