AUTOMATED STOCK TRADING

USING MACHINE LEARNING

GROUP: RISHIKA SHARMA- 210002063
NIRANJANA R NAIR- 210003049

GITHUB REPOSITORY LINK:
https://github.com/gravityinescapable/Stock-Market-Prediction-using-ML

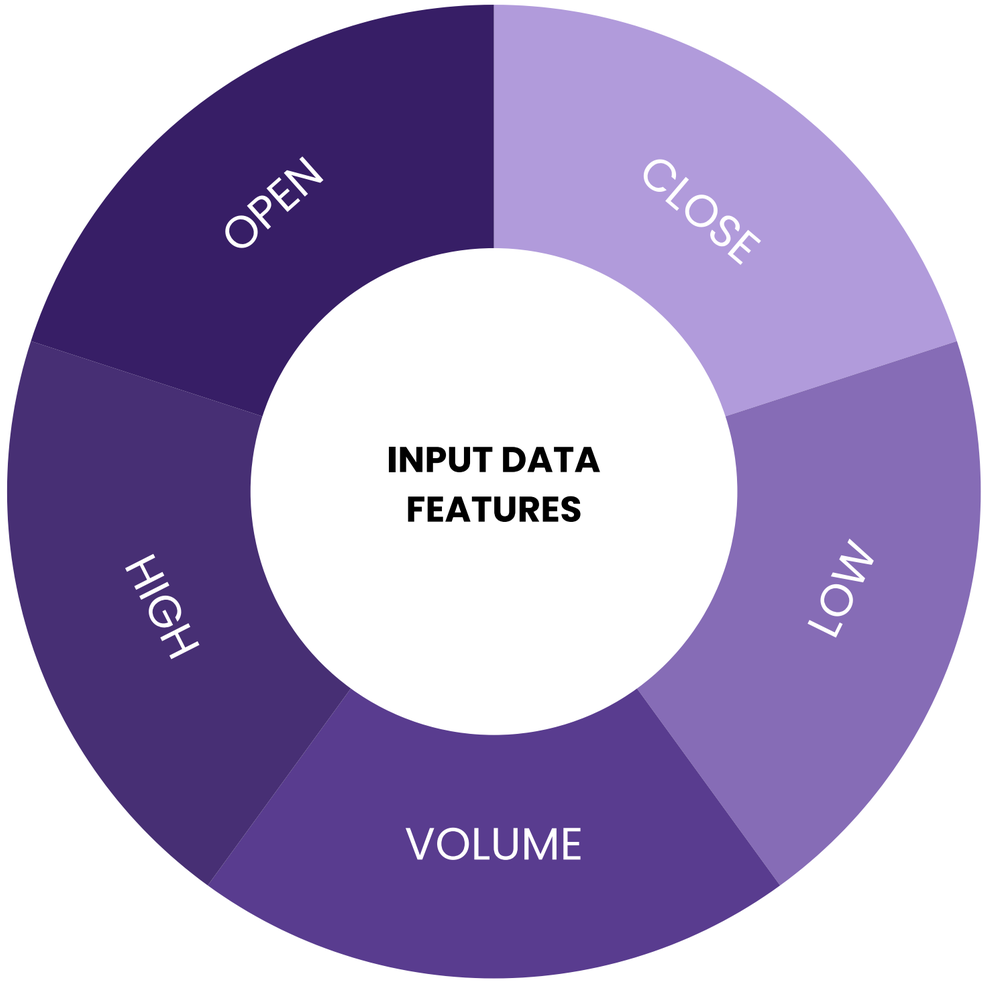
PROBLEM DEFINITION

- Determine whether the price of a derivative would increase or decrease on the basis of current market trends.
- Analyze historical market data and make real-time trading decisions without human intervention.
- Optimize trading strategies and enhance decision-making accuracy in dynamic financial markets.

DATA COLLECTION

- Official NSE data for Nifty 50 index from 2000-2024 (24 years) was selected.

- We define 20 days as one business month- relevance in data splitting and in the calculation of performance metrics.



PREPROCESSING

- Feature Engineering: Created new features from existing data to provide additional information.
- Feature Scaling: Used fit-transform function to standardize the features.
- Data split into training and testing sets

FEATURES

- SIMPLE MOVING AVERAGES
- TRIANGULAR MOVING AVERAGES
- AVERAGE DIRECTIONAL INDEX
- EXPONENTIAL MOVING AVERAGES
- BOLLINGER BANDS
- COMMODITY CHANNEL INDEX
- MOVING AVERAGE CONVERGENCE DIVERGENCE
- KAUFMAN ADAPTIVE MOVING AVERAGES
- STOP AND REVERSE
- MOMENTUM INDICATORS

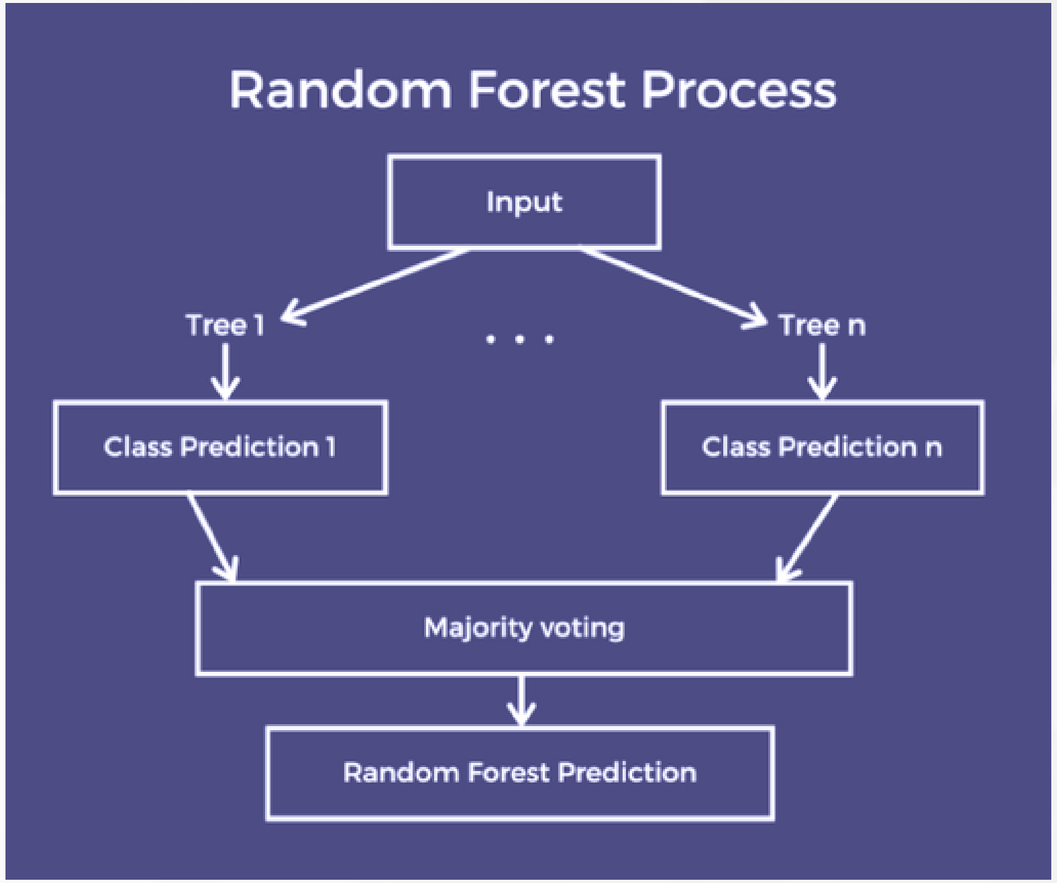
FEATURES

- RATE OF CHANGE
- PERCENTAGE PRICE OSCILLATOR
- RELATIVE STRENGTH INDEX
- STOCHASTIC OSCILLATOR INDICATORS
- ULTIMATE OSCILLATOR
- WILLIAM'S %R
- AVERAGE TRUE RANGE
- TYPICAL PRICE
- VORTEX INDICATORS
- MONEY FLOW VOLUME

ALGORITHMS

- **Random Forest Classifier:** Random Forest combines multiple decision trees offering robust predictions for automated stock trading.

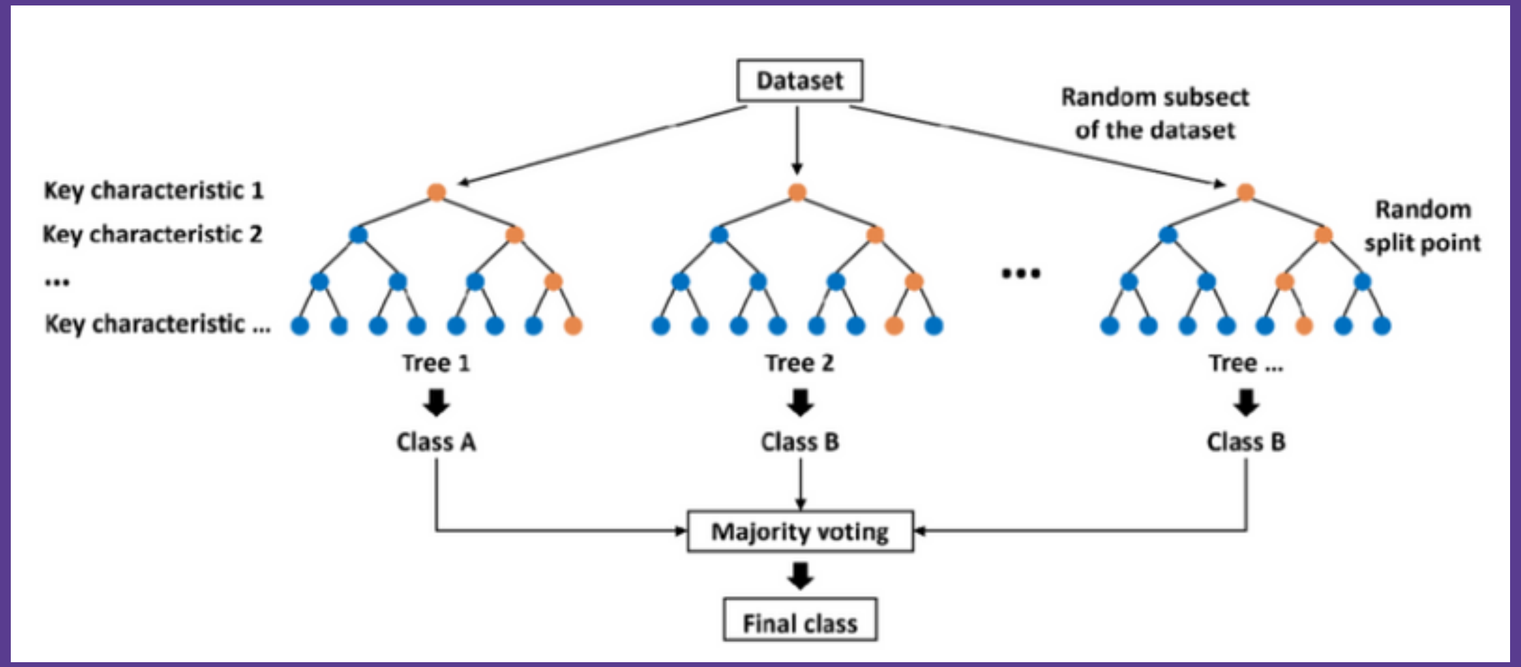
- It provides feature importance scores and handles noisy financial data effectively, making it a popular choice for building automated trading models.



ALGORITHMS

- **Extra Trees Classifier:** Extremely Randomized Trees, injects additional randomness into the tree construction process, reducing overfitting and enhancing generalization.

- Its fast training coupled with feature importance analysis makes it a valuable algorithm for automated stock trading systems.



PERFORMANCE MEASUREMENT

- Several criteria can be employed for performance measurement.
- Provides an objective way to assess the effectiveness of a trading strategy by analyzing various metrics.
- Serves as a feedback mechanism to identify strengths and weaknesses in the approach.

Returns: measure the gain or loss of a trading strategy over a specified period.

Sharpe Ratio: Risk-adjusted return of a trading strategy. Calculated by dividing the excess return by the standard deviation of returns.

Maximum Drawdown: Maximum loss incurred by a strategy from its peak to trough over a specific period.

Calmar ratio: Calculated by dividing the annualized return by the maximum drawdown.

Profit Factor: Measures the ratio of gross profits to gross losses generated by a trading strategy.

Win Loss Ratio: Compares the number of winning trades to the number of losing trades generated by a trading strategy.

TRAINING

- Sliding Window Approach: Utilized a window size of 40 days for training and testing, ensuring fair and accurate fitting of the models.
- Model Selection: Employed Random Forest and Extra Trees Classifier to predict stock labels.
- Lookahead Strategy: Adopted a lookahead of 10 days (2 weeks) for predicting stock label shifts.

TESTING

- Backtesting: Validated ML models by testing trading strategy on historical data.
- Performance Evaluation: Applied strategy on unseen historical data to gauge effectiveness and optimize parameters.
- Risk Assessment: Considered risk-free rate as a benchmark to evaluate risk-adjusted performance.

TRADING STRATEGY

- Utilized Extra Trees Classifier's trained model within our algorithm, processing unlabeled datasets with synthesized features.
- Dynamic Confidence Adjustment: Incorporated a dynamic confidence update mechanism.
- Implemented a rule-based approach to calculate maximum shares tradable per day.

NOVELTY

- **SHAP Integration:** Utilized SHAP to explain model results, enhancing interpretability and understanding of predictions.

- **Dynamic Confidence Factor:** Alpha for confidence update calculated through sentiment analysis of articles taken from web sources.

REFERENCES

- https://towardsdatascience.com/predicting-future-stock-market-trends-with-pytho
- https://towardsdatascience.com/deep-reinforcement-learning-for-automated-stock
- https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForest
- https://scikit-learn.org/stable/modules/generated/sklearn.ensemble. ExtraTreesClas
- https://blog.quantinsti.com/mean-reversion-time-series/
- https://irep.ntu.ac.uk/id/eprint/32787/1/PubSub10294_702a_McGinnity.pdf

THANK YOU