A logo of a company

Description automatically generated A close-up of a logo

Description automatically generated

**Joint Tech Internship Community Program**

**Assignment: Predicting Stock Price Movements in the Stock Market**

**Problem Statement:**

A financial services firm is interested in predicting the future price movements of stocks to aid in making better investment decisions. Your task is to develop a machine learning model that predicts whether a stock's price will go up or down based on historical data and various market indicators.

**Objective:**

Build a predictive model to classify whether a stock's price will increase or decrease the next day. Evaluate the model using appropriate metrics and provide insights into the factors that most influence stock price movements.

**Dataset:**

You are provided with a dataset containing the following columns:

1. **Date**: The date of the stock market data.
2. **StockID**: Unique identifier for each stock.
3. **OpenPrice**: The opening price of the stock on that day.
4. **ClosePrice**: The closing price of the stock on that day.
5. **HighPrice**: The highest price of the stock on that day.
6. **LowPrice**: The lowest price of the stock on that day.
7. **Volume**: The total number of shares traded on that day.
8. **MovingAverage10**: The 10-day moving average of the stock's closing price.
9. **MovingAverage50**: The 50-day moving average of the stock's closing price.
10. **RSI**: The Relative Strength Index, a momentum indicator that measures the magnitude of recent price changes.
11. **MACD**: The Moving Average Convergence Divergence, a trend-following momentum indicator.
12. **Volatility**: The standard deviation of the stock's closing prices over the past 30 days.
13. **Sector**: The sector to which the stock belongs (e.g., Technology, Healthcare, Finance).
14. **PriceMovement**: The target variable indicating whether the stock price increased or decreased compared to the previous day (Up/Down).

**Tasks:**

1. **Data Exploration and Preprocessing:**
   * Load the dataset and perform initial exploration to understand the data.
   * Identify and handle any missing values appropriately.
   * Convert categorical variables into numerical ones using techniques such as One-Hot Encoding.
   * Create additional features if relevant, such as daily price change percentage.
2. **Feature Engineering:**
   * Perform feature scaling (e.g., Standardization or Normalization) on continuous variables.
   * Use dimensionality reduction techniques (e.g., PCA) to reduce the feature space if necessary.
3. **Model Building:**
   * Split the dataset into training and testing sets (e.g., 80-20 split).
   * Train different classification models (e.g., Logistic Regression, Decision Trees, Random Forest, Gradient Boosting).
   * Perform hyperparameter tuning using techniques like Grid Search or Random Search to optimize model performance.
4. **Model Evaluation:**
   * Evaluate your models using appropriate metrics such as Accuracy, Precision, Recall, F1-Score, and AUC-ROC.
   * Compare the performance of different models and select the best one.
   * Analyze feature importance to understand the most significant factors contributing to stock price movements.
5. **Insights and Recommendations:**
   * Provide insights based on your model analysis regarding the factors that influence stock price movements.
   * Suggest actionable strategies for stock trading or investment based on the findings.
6. **Documentation:**
   * Document your process, including data exploration, preprocessing steps, model selection, and evaluation.
   * Include visualizations where necessary to support your findings.

**Deliverables:**

* A Jupyter notebook (or Python script) with the entire workflow.
* A report summarizing your findings, including the model's performance and recommendations for stock trading strategies.

This assignment problem focuses on predicting stock price movements in the financial market, covering various aspects of a typical machine learning workflow, including data preprocessing, feature engineering, model training, and evaluation. It emphasizes the importance of providing actionable insights for making informed investment decisions.

Top of Form

Bottom of Form