

Name : M.Bilal Butt

Intern ID : DHC-356

Task 1: Financial Time-Series Anomaly Detection

1. Dataset Preprocessing Steps

- Downloaded Apple stock data (AAPL) from Yahoo Finance (2020-2025).
- Flattened MultiIndex columns to simplify analysis.
- Identified the 'Close' column dynamically to handle naming variations.
- Calculated 20-day SMA and Bollinger Bands (Upper and Lower).
- Dropped NaN values after rolling calculations.

2. Model Selection and Rationale

- Isolation Forest was chosen due to its robustness, ease of use, and suitability for unsupervised anomaly detection on continuous data.

3. Challenges Faced and Solutions

- **MultiIndex columns:** Flattened using `data.columns = [' '.join(col).strip() for col in data.columns.values]`.
- **Missing Data:** Dropped rows only after indicator calculations.
- **Dynamic column naming:** Handled dynamically using a search on 'Close' substring.

4. Results with Visualizations and Interpretations

- Detected approximately 5% anomalies.
- Anomalies mostly correspond to sudden price spikes or drops.
- Visualized using both Matplotlib (static) and Plotly (interactive).

Screenshots

