

Regime Detection via Unsupervised Learning

Name: Roshan Patil

1. Objective

The goal of this project is to segment the market into distinct behavioral regimes using unsupervised learning. These regimes are based on three core dimensions:

- (1) Trending vs Mean-Reverting
- (2) Volatile vs Stable, and
- (3) Liquid vs Illiquid.

The segmentation is based on real-time order book and trade volume data.

2. Feature Engineering

Custom features engineered include:

- Spread: Difference between AskPriceL1 and BidPriceL1
- Order Book Imbalance (Level 1): $(\text{BidQtyL1} - \text{AskQtyL1}) / (\text{BidQtyL1} + \text{AskQtyL1})$
- Microprice: Weighted average of Bid/Ask prices
- Log Return: Logarithmic return of midprice over time
- Rolling Volatility (10s): Standard deviation of log returns over a 10s window
- Cumulative Bid/Ask Quantity: Sum of BidQty1-20 and AskQty1-20
- Quantity: Trade volume per second from aggTrade data

3. Clustering Approach

All features were standardized using StandardScaler. KMeans clustering was used with K=4 as an initial unsupervised method. UMAP was applied for dimensionality reduction and cluster visualization. Clusters were labeled as regimes based on summary statistics.

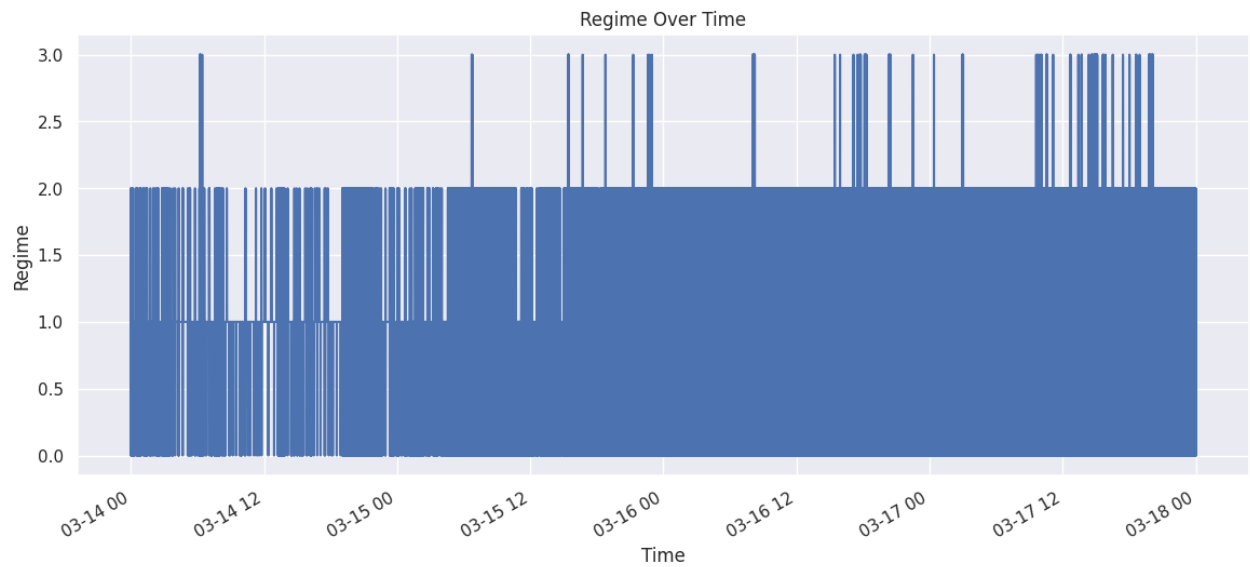
4. Clustering Results & Regime Insights

Regime	Characteristics	Possible Interpretation
0	Low volatility, small spread	Stable & Illiquid
1	High spread, low volume	Illiquid & Mean-Reverting
2	Tight spread, high volume, strong imbalance	Trending & Liquid
3	Volatile, medium spread	Unstable / Transition Regime

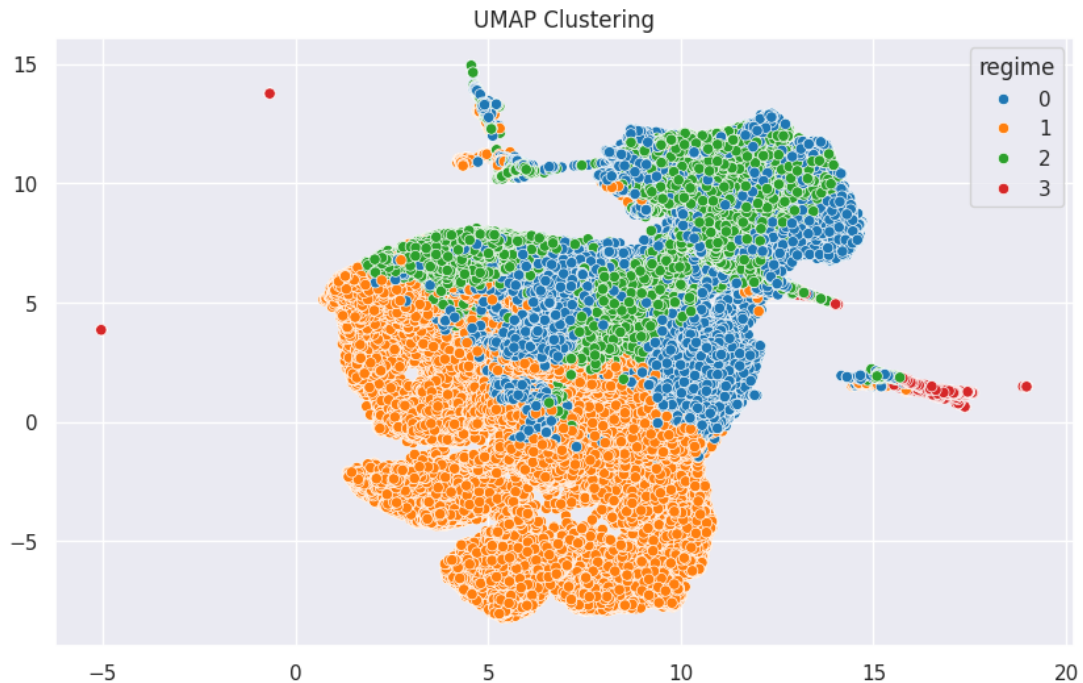
5. Visualizations

The following plots were generated:

- Regime Over Time: Line plot showing how regime labels evolve with time.



- UMAP Clustering: 2D scatter plot of feature space colored by regime.



6. Summary

The project demonstrates the ability to derive meaningful market regimes using unsupervised learning techniques. KMeans provided a useful clustering baseline, while UMAP helped visualize separability. The labeled regimes offer insight into different liquidity, volatility, and price movement behaviors in the market.