

Detecting Financial Market Regimes

Blake Krouth
College of Computing
Michigan Technological University
Houghton, MI, USA
bjkrouth@mtu.edu

Nandhika Rajmanikandan
College of Computing
Michigan Technological University
Houghton, MI, USA
nrajmani@mtu.edu

Ganesh Vannam
College of Computing
Michigan Technological University
Houghton, MI, USA
gvannam@mtu.edu

Abstract—This report explores various feature-engineering and clustering techniques on time-series S&P 500 stock market data with the goal of detecting market regimes, or periods of similar behavior. The models were evaluated with both objective clustering metrics and the subjective measure of interpretability. The hidden Markov model was found to perform the best with simple, noise-free feature sets. K-means and fuzzy c-means performed best with more complex feature sets, tending to produce simple and easily interpretable clusters. The Gaussian mixture model did not produce high-confidence results, but visually produced smooth regime switches when fit to a feature set containing long-term stock trends. Without the use of a sliding window or out-of-sample testing, the results are limited in scope but serve as a starting point for a robust market forecasting or historical market visualization system.

Index Terms—finance, clustering, data mining, machine learning

I. INTRODUCTION

Understanding the trend of the financial market is vital to financial investors who wish to intelligently manage their portfolios and assets. Due to the unobservable and complex workings of the financial market, market trend detection is not an easy task. Detecting market regimes is one approach to this problem. In the context of a financial market, a regime refers to a period of similar behavior in the market [1], [2]. By understanding the regime, or overall behavior, of a financial market, an investor can create a better strategy. The goal of this report is to explore various clustering and feature engineering techniques in order to create clusters that accurately and intuitively describe market regimes in way that would be of use to investors.

II. RELATED WORKS

Recent attempts to detect market regimes rely on unsupervised learning techniques that assume latent variables (e.g., market returns or macroeconomic indicators) determine the market state. In [3], fuzzy C-Means clustering was applied to a sliding window over multiple time-series data of both synthetic and real-world economic indicators. While the authors were more interested in clustering as a precursor to market forecasting, they were successfully able to create clusters that were fed into their forecasting model. In [4], the financial metric of monthly realized covariances is used to create both statistical (i.e., TVAR, LSTVAR, and MSVAR) and clustering (i.e., AGNES) models. In model evaluation on a synthetic

dataset, the clustering model performed the best, and was considered by the authors to be the best performer alongside the LSTVAR statistical model. [2] take a unique approach by first training a k-means clustering model on Federal Reserve Economic Data (FRED) to cluster market regimes and next training various supervised models on both the FRED data and clustered regimes to classify out-of-sample data.

[1] propose a framework for comparing unsupervised models in the context of market regime detection. The framework addresses common problems in market regime detection, such as models not adapting to changing market conditions, ensuring consistent labeling as regimes change over time, and choosing the appropriate number of regimes. Notably, the authors opt for a rolling window retraining, similar to what was done in [3]. Using their framework, [1] evaluated a Hidden Markov Model with Gaussian observation model and a Hidden Markov Model with Gaussian mixture observation model that was fitted to FRED indicators. The process was able to generalize well to other macroeconomic indicators, such as futures and mutual funds.

A frequent drawback of market regime detection models is that they assume a fixed number of regimes that are chosen during training and are unable to adapt to dynamic market conditions in which the number of regimes vary [1], [3]. [3] address this by dividing the data with a sliding window that dynamically chooses the appropriate number of regimes, while [1] describes a thresholding policy. However, [1] only addresses the condition that an additional regime should be added, not that the number of regimes could decrease, which they state is logical since a regime condition does not disappear and there should be historical consistency in the model.

A theme common to [3] and [4] is the use of a synthetic dataset with known, generated regimes to perform a quantitative evaluation on clustering performance. While potentially ideal for model evaluation, synthetic dataset generation is out of the scope of the class and our goals.

With the exception of [1], the reviewed literature fails to illustrate the trade-offs of various clustering techniques when fit to financial data. Additionally, the time-series data used tends to be general, indicating the trend of the market as a whole instead of a specific ticker. We address these shortcomings by evaluating the tradeoffs of the models used and applying them to engineered features from a specific stock.

III. METHODOLOGY

A. Data Acquisition, Preprocessing, and Feature Engineering

Historical daily time-series stock data of the S&P 500 (ticker ^GSPC) from January 1, 1990 to January 1, 2024 was obtained from Yahoo Finance [5], resulting in 8565 ratio observations containing the daily close, high, low, open, and volume prices.

Using the data from [5], the daily return, (1), was calculated using the day's closing price, p_t and the previous day's closing price p_{t-1} .

$$r_t = \frac{p_t}{p_{t-1}} - 1 \quad (1)$$

Additionally, the momentum, or percent change in return, over a 10-day period was calculated, and the volatility, or standard deviation, of the daily returns was calculated on a 10-day rolling basis.

Due to the noise present in daily time-series financial data, this data was then resampled monthly, with the last value of each month retained. The monthly return was calculated according to (1) using the month's closing price, p_t and the previous month's closing price p_{t-1} .

The compounding cumulative return, (2), was then calculated for each month using the monthly returns up until that point in time.

$$c_t = \prod_{i=1}^t (1 + r_i) \quad (2)$$

Finally, the volatility, or standard deviation of the daily returns was calculated on both a 10-day (2 weeks) and 21-day (1 month) rolling basis. Once again, the motivation behind the rolling basis was to reduce the noise present in the financial data.

At this point, all NA values were dropped to account for the rolling calculations, resulting in the monthly returns that can be seen in “Fig. 1.”

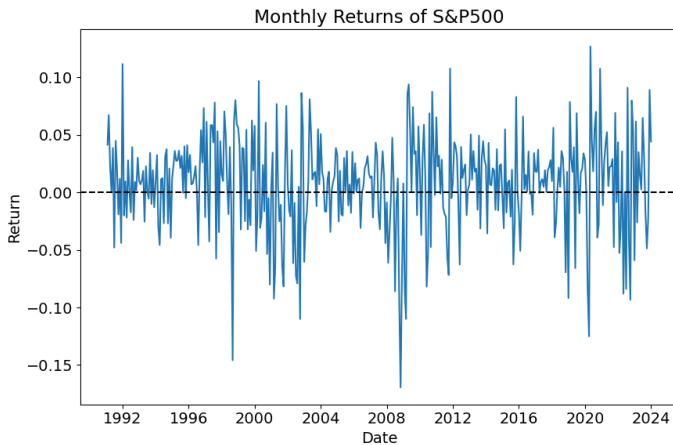


Fig. 1. Monthly returns of S&P 500.

While the monthly returns are visibly noisy, the distribution of the monthly returns appeared to be roughly normal without significant outliers, as seen in “Fig. 2,” indicating that efforts were successful in removing destructive noise from the data.

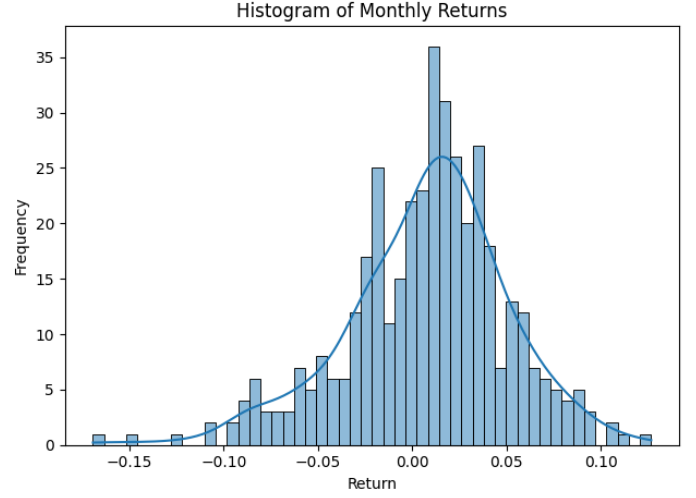


Fig. 2. Distribution of S&P 500 monthly returns.

However, additional features with further reduced noise were also created: the 6 month moving average of the monthly cumulative return, the 6 month moving average of the monthly volatility, the 3-month percent change (momentum) in monthly closing price, and a z-score representing the month's return in comparison to the rolling average for the monthly returns in the past 12 months.

EDA was performed on these 4 features to discover their distributions and ensure noise was minimal. The 6-month cumulative returns were right skewed, the 6-month rolling volatility was heavily right skewed, the 3-month momentum was relatively normally distributed, and the yearly z-score was slightly left skewed.

In total, 10 additional features were engineered:

- The daily return to capture short term trends
- The 10-day percent change in returns (momentum) to capture medium-term trends
- The 10-day volatility in returns to capture medium-term stability
- The month-to-month return to capture monthly return trends
- The cumulative monthly return to capture long-term monthly changes
- The rolling monthly volatility to capture month-to-month stability
- The 6 month moving average of the monthly cumulative return to reduce noise
- The 6 month moving average of the monthly volatility to reduce noise
- The 3-month percent change in monthly closing price
- A z-score representing the month's return in comparison to the rolling average for the monthly returns in the past

12 months

B. Clustering with Simple (Univariate) Feature Set

The clustering techniques chosen were k-means, the hidden Markov model (HMM), fuzzy c-means (FCM), and the Gaussian mixture model (GMM). For a baseline, these models were first evaluated with only one feature, the 6-month rolling average of cumulative returns, which provided insight into how the models behaved. The appropriate number of regimes (clusters/hidden states) was obtained through hyperparameter tuning with clustering performance metrics. The results are summarized in “Table I.”

TABLE I
CLUSTERING METRICS FOR UNIVARIATE FEATURE SET

Model	Number of Regimes	Performance Metric
k-means	5	silhouette score = 0.66
HMM	4	log-likelihood = 58.74
FCM	3	FPC = 0.8339
GMM	5	log-likelihood = -370.53

C. Clustering with Multivariate Smoothed Feature Set

The models were next evaluated with three smoothed features, summarized in “Table II.” The goal with this feature set was to capture several long-term patterns in the data.

- The rolling monthly volatility
- The 3-month percent change in monthly closing price
- The z-score representing the month’s return in comparison to the rolling average for the monthly returns in the past 12 months

TABLE II
CLUSTERING METRICS FOR MULTIVARIATE SMOOTHED FEATURE SET

Model	Number of Regimes	Performance Metric
k-means	3	silhouette score = 0.35
HMM	6	log-likelihood = -1042.51
FCM	3	FPC = 0.5261
GMM	4	log-likelihood = -1420.10

D. Clustering with Multivariate High Frequency Feature Set

Finally, the models were evaluated with three daily features, summarized in “Table III.” This feature set was created to capture short-to-medium-term trends.

- The day-to-day return
- The rolling 10-day volatility
- The 10 day percent change (momentum)

IV. DISCUSSION OF RESULTS

A. Clustering with Simple (Univariate) Feature Set

In general, the results from the univariate feature set were poor. Even more, the identified clusters spanned several years, likely due to the stability of the 6-month smoothed cumulative return that it was fit to, making them only useful in a historical sense. However, using less-smoothed features, like the monthly

TABLE III
CLUSTERING METRICS FOR MULTIVARIATE HIGH FREQUENCY FEATURE SET

Model	Number of Regimes	Performance Metric
k-means	2	silhouette score = 0.411
HMM	6	log-likelihood = -18455.86
FCM	2	FPC = 0.6603
GMM	7	log-likelihood = -27872.36

return, had the opposite effect. The results were extremely noisy and hard to interpret.

The k-means and fuzzy c-means detected very similar clusters with strong cluster separation. Visually, the clusters appeared to capture general transitions of market trends instead of sharp transitions that one may expect, like a transition from bullish to bearish, for example. The Gaussian mixture model was the lowest performer, producing clusters that were visually ambiguous with low confidence. The hidden Markov model produced the most easily interpretable clusters, as seen in “Fig. 3.”

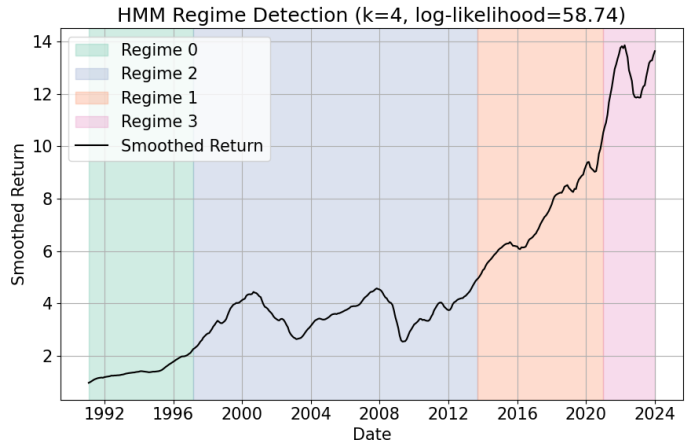


Fig. 3. HMM Market Regimes Visualised Over 6-Month Smoothed Cumulative Returns.

One may intuitively describe the clusters as

- Cluster 0: Moderate but consistent bullish trend
- Cluster 1: Volatile period
- Cluster 2: Very bullish period
- Cluster 3: Strong bullish period with high volatility

Here, the hidden Markov model captures temporal transitions the best. This is intuitive, since other clustering methods used, like k-means for example, simply create spherical clusters that group data based on similar feature sets with no time-series context, while the hidden Markov model captures time-series context with its concept of probabilistic regime changes that depend on previous states.

While using one feature produces acceptable clustering metric scores, it was apparent that the clusters were hard to interpret, with the exception of those produced by the HMM. Furthermore, with data as complex as financial data, it seems unlikely that trends can be accurately described with only

one feature. This motivated further exploration of the two multivariate feature sets.

B. Clustering with Multivariate Smoothed Feature Set

Interestingly, the multivariate smoothed feature set objectively had significantly worse clustering results than the univariate feature set according to clustering evaluation metrics. However, as seen in “Fig. 4,” the results appear to be easier to interpret. While clusters are necessarily ambiguous, one may describe the regimes identified by the models as follows:

For k-means,

- Regime 0: A downward but rebounding trend
- Regime 1: A downward trend
- Regime 2: A positive trend

For the HMM, the larger number of clusters, 6, makes visual analysis more difficult. Even more, the regimes are much noisier, with the exception of regime 2, which seems to identify a non-volatile upward trend. In this case, the poor log-likelihood score seems to align with poor interpretability.

For the FCM, regime 2 is the most unique, appearing to capture market transitions instead of the more polar upward/downward trends identified by k-means, for example. FCM’s soft boundaries differentiate it from k-means here, which was notably not the case for the univariate feature set.

- Regime 0: long-term upward trends
- Regime 1: local downward trends
- Regime 2: market peaks/dip transitions

Once again, the GMM performed poorly, reporting low confidence and visually ambiguous clusters.

- Regime 0/1: Market transitions
- Regime 2: General upward trend
- Regime 3: Ambiguous but corresponds with a single sharp dip

While these are all human-biased and necessarily incorrect assumptions due to the nature of clustering, there is certainly a general trend.

C. Clustering with Multivariate High Frequency Feature Set

Like the multivariate smoothed feature set, the clustering performance metrics for the multivariate high-frequency feature set were significantly worse than the univariate feature set. While expectedly noisier than the smoothed feature set, the clusters are once again easier to interpret than those identified by the univariate feature set, as seen in “Fig. 5.”

For both k-means and FCM two regimes were identified, which were subjectively the easiest to interpret out of all clusters. FCM appeared to identify more downward trends than k-means.

- Regime 0: Downward trend
- Regime 1: Upward trend

As was the case with the multivariate smoothed feature set, the HMM and GMM regimes are difficult to interpret visually and have low confidence.

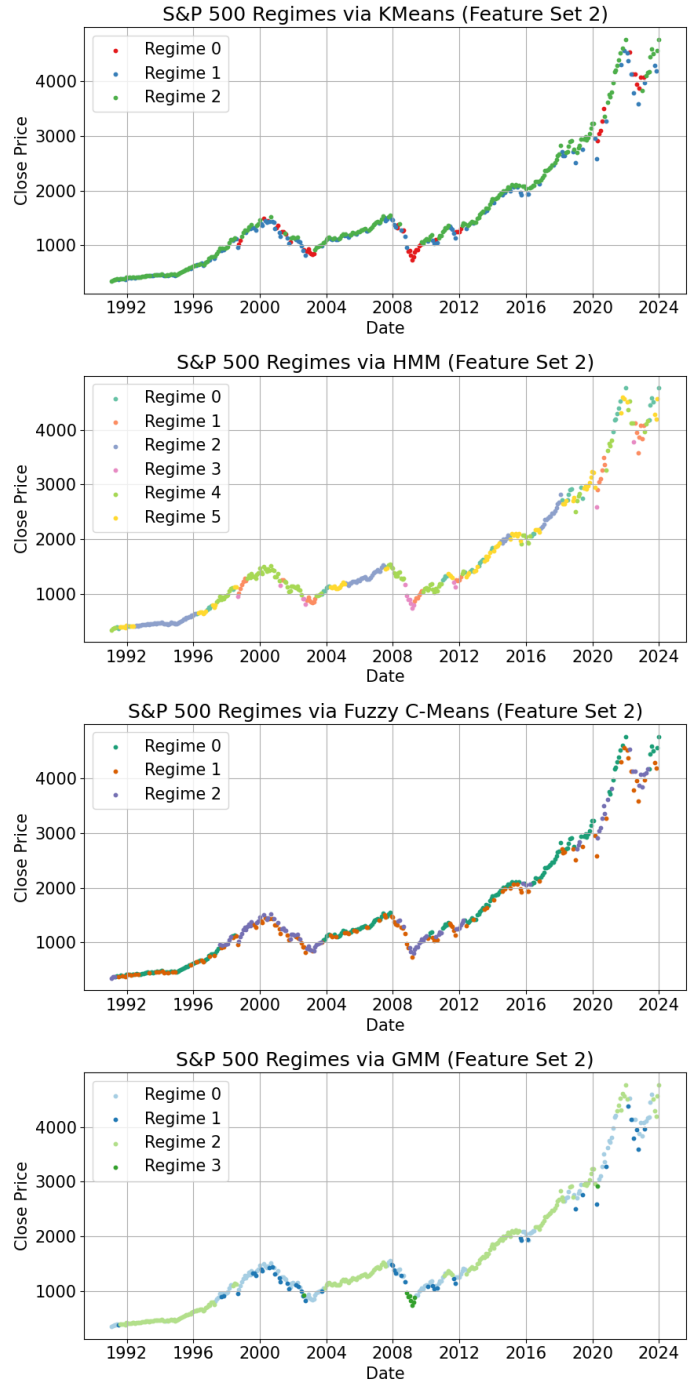


Fig. 4. Model Regimes Visualized Over Close Price for Multivariate Smoothed Feature Set.

V. CONCLUSION

The goal of this report was to identify periods of similar behavior, or regimes, in time-series stock market data using various clustering techniques and engineered feature sets. This was achieved by fitting k-means, fuzzy c-means, hidden Markov, and Gaussian mixture models to 3 feature sets.

- 1) A univariate feature set created as a baseline that captured long-term stock trends

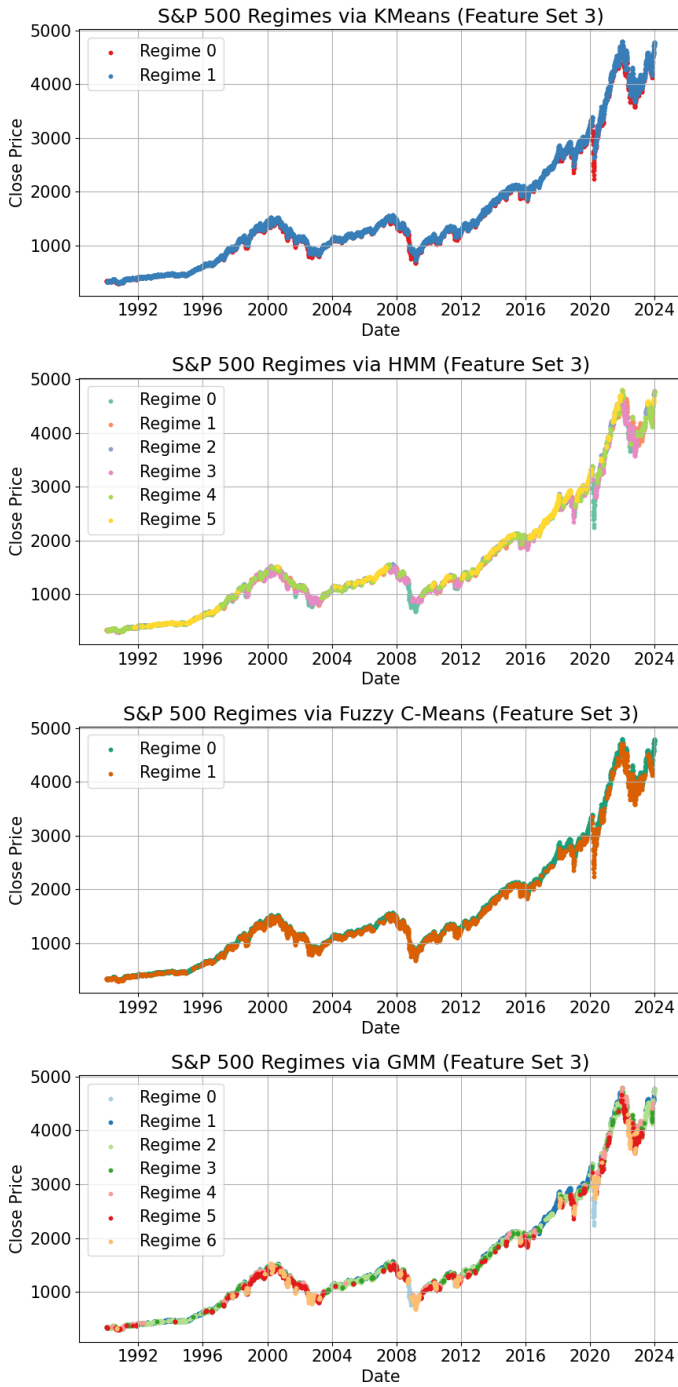


Fig. 5. Model Regimes Visualized Over Close Price for Multivariate High Frequency Feature Set.

- 2) A multivariate smoothed feature set that captured more complex long-term stock trends
- 3) A multivariate high-frequency feature set that captured more complex short-to-long-term stock trends

From the univariate feature set, it was observed that the hidden Markov model was the best performer with limited features, while other models failed to produce visually meaningful and unambiguous clusters. More complex trends were explored in

the second and third feature sets. Notably, the hidden Markov model tended to produce complex and visually ambiguous clusters in higher feature spaces. The GMM in general was a low performer but did provide interpretability with the second feature set, capturing long-term, smooth regime switches. K-means and fuzzy c-means produced easily interpretable but potentially oversimplified clusters in both multivariate feature sets.

One major limitation of this report's work is that it assumes a fixed time frame (i.e., January 1, 1990, to January 1, 2024). The work of this report could be built upon by incorporating a sliding window technique as done in [1], [3] to ensure that the time-frame of the stock data does not affect the clustering results and that regimes are consistently labeled over time. Additionally, out-of-sample testing was not done, so it cannot be said with confidence that the feature-engineering techniques are universal to other stock data. Further testing is necessary.

While limited in scope, the results of this report do provide insight into the significant effects that feature engineering has on time-series financial data and the conditions in which certain clustering techniques have the potential to succeed. This work, with the mentioned improvements, could be used as the basis for a robust forecasting system, like the one proposed in [3], or as a visualization tool for investors to analyze historical stock trends.

REFERENCES

- [1] A. Hirsra, S. Xu, and S. Malhotra, "Robust rolling regime detection (r2-rd): A data-driven perspective of financial markets," To be published. Accessed: Mar. 1, 2025. [Online]. Available at SSRN.
- [2] P. Akioyamen, Y. Z. Tang, and H. Hussien, "A hybrid learning approach to detecting regime switches in financial markets," presented at ACM International Conference on AI in Finance, New York, NY, USA, Oct. 15-16, 2020. doi: <https://doi.org/10.1145/3383455.3422521>
- [3] R. Chen, M. Sun, K. Xu, J. Patenaude, and S. Wang, "Clustering-based cross-sectional regime identification for financial market forecasting," in *Database and Expert Systems Applications*, vol. 13427, pp. 3-16, 2022. doi: https://doi.org/10.1007/978-3-031-12426-6_1
- [4] A. Bucci and V. Ciciretti, "Market regime detection via realized covariances," presented at 14th International Conference on Computational and Financial Econometrics, virtual, 2020. doi: <https://doi.org/10.1016/j.econmod.2022.105832>
- [5] Yahoo Finance, "S&P 500 (^GSPC) Stock Price" Yahoo Finance. [Online]. Available: <https://finance.yahoo.com/quote/%5EGSPC/history/>

VI. SELF-DECLARATION

- 1) Blake Krouth: Data acquisition, EDA, and the univariate analysis. I conducted the literature review and wrote all text in this report.
- 2) Ganesh Vannam: GMM and K-Means clustering models, created visualizations for model interpretation, and tested the models across different time frames to assess performance consistency.
- 3) Nandhika Rajmanikandan: Led the experimentation with different feature sets, introducing the idea of comparing single vs. multivariate inputs to assess model performance. I implemented the Fuzzy C-Means and HMM clustering models with appropriate evaluation metrics and explored GARCH modeling and change point detection as potential extensions beyond the current report.