Proceedings of the 4th International Conference on Applied Physics and Mathematical Modeling
DOI: 10.54254/2753-8818/2026.37205

# *Unsupervised Learning for Financial Market Regime Detection Using Volatility Features*

**Ruyue Xu**

*Shanghai Pinghe School, Shanghai, China*
*jiuyue733@gmail.com*

***Abstract.*** Financial markets fluctuate under several relatively persistent conditions, and returns, volatility, correlation, and risk patterns vary depending on the conditions. Because market regimes and their boundaries are not directly observable and widely recognized labels are unavailable, unsupervised learning provides a suitable framework for identifying hidden states from financial time series. This review examines the transformation of volatility-related features into model inputs for clustering and latent state detection. In order to illustrate the challenges posed by non-stationarity, outliers, and unequal feature scales, it first explains core variables including log returns, rolling volatility, drawdown, trading volume, implied volatility, and dynamic correlation. Then, it compares K-means, hierarchical clustering, Gaussian mixture models, spectral clustering, and hidden Markov models in terms of assumptions, temporal coherence, probabilistic output, interpretability, and evaluation. According to recent studies, market regime identification approaches are evolving from single-asset, hard-label classifications to multidimensional, probabilistic, rolling, and information-enriched systems. Nevertheless, the results remain sensitive to feature design, window length, state number, initialization, sample period, and validation criteria. Future research should focus on online updating, uncertainty quantification, cross-market robustness, and economically meaningful interpretation.

***Keywords:*** market regime detection, volatility features, unsupervised learning, clustering

## 1. Introduction

Financial Markets are not always under a single stable statistical system. Certain events, such as corrections, turbulent recoveries, prolonged declines or crisis periods in which volatility and cross-asset correlations rise simultaneously, may occur after gradual appreciation and low volatility. A Market Regime can be regarded as a relatively stable market state. At this time, returns, risks, dependencies and investor behaviour all have similar characteristics. Regime-switching research is necessary because sudden changes in means, variances, autocorrelations and cross-covariances can explain several typical features of financial data, such as heavy tails, volatility clustering, skewness and time-varying correlations [1]. The new system will change the diversification of the portfolio and risk estimation for trading and forecasting models.

The first is that the market system has not been released and cannot be observed directly. Prices, returns, option-implied volatility, trading volume and macroeconomic indicators are all observable,

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/2026.37205

but there is no official daily index for the bull, bear, calm or crisis state. Classical regime-switching models generally assume that the data are generated by an unobserved state that changes over time. Hamilton's Markov-switching model is a general form in which the parameters change according to a discrete-state Markov process, and researchers can probabilistically infer the hidden state by observing the time-series behaviour [2]. Another way in modern data science is clustering; that is, groups of observations or time windows are formed based on their similarity without using pre-defined class labels.

Regime descriptions are also convenient if they will not be used as direct forecasts. They can show whether the risk model was only tested in normal circumstances, whether diversification reduced the risk, and whether the performance remained stable during a crisis. Thus, market regime analysis seeks to describe and diagnose the market. The utility of a cluster does not depend on whether it has an attractive name; rather, it should have different, stable and repeatable returns and risk characteristics.

Volatility features are used to detect regime shifts, and changes in market risk are generally more stable and easily identified than changes in average returns. A regime detector can use daily log returns, rolling standard deviations, realised or implied volatility, maximum drawdown, trading volume and dynamic correlations. These variables are the numerical representations of economic concepts such as market calm, stress, contagion and recovery for the model. However, this feature construction is not fair. A five-day volatility window responds quickly but has considerable noise; a sixty-day window is smoother but may not show a crisis in time. However, a distance-based algorithm may be dominated by the largest-valued feature if raw variables with different scales are used.

Now, research on market regime detection has moved away from fixed thresholds and single-variable volatility models. Structural clustering combines change-point and distribution tests to find repeated volatility patterns [3]. Wasserstein K-means methods have been extended to multidimensional financial series, and clustering is performed by comparing empirical distributions over time-series windows [4]. Other studies have added correlations, macroeconomic conditions, probabilistic state assignments, online retraining, high-frequency data, option-implied distributions and news-derived information. This development has connected financial reasoning with statistical feature engineering and unsupervised learning.

The three problems addressed here are: First, how can financial time series be transformed into informative volatility-related features? Second, what are the differences in assumptions, outputs, strengths and weaknesses of the main unsupervised learning methods? Third, why are papers discussing the same market often classified under different regimes? The above paper will introduce the general ideas of market regimes and volatility features, compare the main detection methods, summarize some representative recent research, and finally discuss the shortcomings and future directions. The aim is not to reproduce the difficult experiments or derive all the algorithms, but to explain why a cluster may or may not be considered an economically significant market regime.

## 2. Basic concepts of financial market regimes and volatility features

The beginning of most market regime studies is usually a series of asset prices. As prices generally have a trend and are not directly comparable across different assets, researchers often use simple returns (which are the percentage changes in price) or log returns (which are logarithmic changes in price) instead. Because they are cumulative over time and nearly the same as simple returns for small changes, log returns are convenient. The sign of the return indicates the direction, and its

DOI: 10.54254/2753-8818/2026.37205

absolute value is the size of the price change. The return alone cannot tell us whether the market has shown a stable trend or fluctuations and ups and downs.

Volatility generally refers to the standard deviation of returns over a rolling window, and it shows the dispersion of returns. For each day, a window of the previous observations is shifted forward, and then the statistic is recalculated. Short windows are highly responsive but unstable; long windows reduce noise but may smooth out rapid changes. Historical or realised volatility is the fluctuation of the market in the past, and implied-volatility indices, such as the Cboe Volatility Index (VIX), reflect the market's expectation of future volatility. To quantify the total downside risk that standard deviation does not cover, maximum drawdown measures the drop from a previous high. Trading volume shows changes in the market, and rolling correlations show changes in asset correlation. Under stress, correlations may rise and thus reduce the protective effect of diversification [1].

Financial time series present additional problems for machine learning. The distributions are non-stationary, and thus the parameters estimated in one period may not be applicable to others. Extreme observations may be data errors, but they can also be crisis events that the model is supposed to identify; therefore, they should not be automatically excluded. Market holidays, asynchronous trading calendars, or the first observations of a rolling calculation may all result in missing values. Returns, VIX levels and volume are also different in terms of units and numerical ranges. Distance directly may cause large-scale variables to dominate the clustering results. Therefore, features are generally normalised before clustering to make them more comparable.

The above indicators can form a feature vector that describes a particular day or time period. For example, a period of high returns and low volatility may be a stable uptrend; a continuous decline in returns and drawdowns indicates a recession; and high volatility, a rising VIX, and significant drawdowns all suggest market turmoil. High volatility is not necessarily a decrease; therefore, both the direction and the risk of returns should be considered. As the market regime does not have consistent and directly observable ground-truth labels, unsupervised methods such as clustering can be employed to explore latent regimes based on feature similarity.

## 3. Unsupervised learning methods for financial market regime detection

### 3.1. Data preparation and feature representation

Align the trading dates, handle missing values, check for recording errors, construct rolling variables and standardise the features before clustering. A feature vector of returns, volatility, VIX, maximum drawdown and changes in trading volume can be used to represent each day. Clustering individual daily observations may result in frequent regime switches; therefore, researchers can divide the time series into windows and examine the distributions of returns and volatility in these windows instead. This structure clustering further divides a non-stationary series into locally stationary parts and groups similar parts as belonging to the same market regime [3].

### 3.2. K-means clustering

K-means clustering divides the data into K clusters by minimising the distance between each data point and the centre of its cluster. K-means has a small amount of computation and is simple to understand. A cluster of negative returns, high volatility and a large drawdown may be due to market pressure; on the other hand, a cluster of positive returns and low volatility can be regarded as a stable uptrend. The model has some deficiencies; for example, the number of clusters must be set

DOI: 10.54254/2753-8818/2026.37205

beforehand by the researcher, and the results are easily affected by outliers, feature scaling and the initial cluster centre.

Some extensions choose to compare the probability distributions of time windows rather than individual data points. Distributional and covariance-based methods can handle the differences in returns, volatility and correlation of individual observations by considering distributions [4, 5]. The selection of the window length and distance measure will still affect the final results.

### 3.3. Hierarchical and spectral clustering

Hierarchical clustering starts by considering all observations as individual clusters and then gradually merges the most similar clusters. A dendrogram can be used to find a reasonable number of clusters and the relationship among the regimes. However, the result will be affected by the chosen distance and linkage method, and incorrect early merges cannot usually be corrected.

Build a graph based on the similarity of observations, and then spectral clustering can find non-spherical clusters that K-means fails to detect. Researchers have previously combined change-point detection, distributional distance and clustering to find volatility regimes in many financial assets [3]. However, the selection of parameters and economic interpretation are generally less intuitive than those for K-means.

### 3.4. Gaussian mixture models

A Gaussian Mixture Model assumes that the data are generated from several Gaussian distributions with different weights. It can offer probabilistic membership under multiple regimes for each observation, and is not a hard assignment as in K-means. Therefore, it is more suitable for a market with an undefined boundary and slow changes. Financial returns are generally heavy-tailed and do not meet the Gaussian assumption. Researchers also need to consider the risk of unstable estimation and determine how many components and what covariance structure to use.

### 3.5. Hidden Markov models and temporal coherence

A Hidden Markov Model (HMM) assumes that the market environment is an unobservable state that changes over time according to transition probabilities. Even if there is a short-term fluctuation in the data on a particular day, the model may not recognise an immediate regime change because it can also recognise the continuity of market conditions [2]. Therefore, although an HMM is not a typical clustering method, it can infer latent market states without labelled data.

The results of an HMM depend on the number of states, starting values and distributional assumptions. Jump models improve temporal consistency by penalising frequent state changes, and rolling estimation updates regime estimates with the addition of new observations [6, 7].

### 3.6. Evaluation

As the market regime does not provide consistent ground-truth labels, only classification accuracy cannot be used to assess the model. Silhouette Score measures the degree of cluster cohesion and separation, and a higher value is generally better. Cluster dispersion and overlap are shown by the Davies-Bouldin Index, and a smaller value is better. Determine if the market regime has been stable over time and whether the results change after modifying features, window length or the sample period. Finally, the regimes need to have economic reasons. For example, a high-volatility period is

DOI: 10.54254/2753-8818/2026.37205

generally associated with a high VIX or a large drop. Therefore, a good model should have economic interpretability, stability, temporal consistency and separate clusters simultaneously.

## 4. Recent studies on unsupervised financial market regime detection

Recently, studies have shown that the results of market regime identification are very sensitive to both the data used and the information in the feature space. Luan and Hamp have both studied synthetic data and real multi-dimensional data, and the results show that sliced Wasserstein K-means can classify market regimes by comparing window distributions instead of individual observations [4]. Bucci and Ciciretti use realised covariance matrices and hierarchical clustering to identify regimes of changing cross-asset dependence [5]. If the change in correlation is of the same order as the change in individual volatility, then the above method can be used.

Macroeconomic Data are the overall conditions of the economy. Oliveira and others have taken a large macroeconomic dataset and used a modified K-means method to maintain the consistency of classification over time. They extend deterministic labels to regime distributions and connect market condition forecasts with tactical asset allocation [8]. The two changes in the literature can be summarised as follows: first, regime detection has moved from single return series to high-dimensional information; second, researchers are increasingly representing uncertainty in current or future states rather than relying solely on hard labels.

Temporally stable is now a primary problem. Aydinhan and others refer to regimes as continuous periods with similar behaviour and employ a continuous statistical jump model to estimate the probabilities of different conditions [6]. Their mode-loss mechanism promotes sharper regime assignments and keeps probabilistic transitions. Kim and Kwon evaluate a regime-based dynamic asset allocation strategy out of sample and account for transaction costs [9]. The model of the judging regime should also be robust and have practical risk consequences, not only fit the in-sample data. Pagliaro addresses temporal updating with a rolling HMM that is refitted using only past observations [7]. Thus, regime detection can be updated over time without the need for future information.

Other research combines latent-state models and prediction systems. Koukorinis et al. combine HMM-generated features with support vector and multiple-kernel machines for high-frequency financial regime classification [10]. A generative temporal model can be used to summarise microstructural changes and then a discriminative model performs classification. Although this hybrid system is more complicated than the methods required for a simple review, it shows that an HMM can be used as a generative feature extractor in a discriminative classification pipeline.

The selected asset class may also change the meaning of a regime. Ardia et al. apply Markov-switching generalised autoregressive conditional heteroskedasticity (GARCH) models to Bitcoin returns and identify high-volatility and low-volatility regimes [11]. In that study, regime-switching models performed better than single-regime models in the estimation of Value-at-Risk. Therefore, a volatility classification can also provide economic information based on different features than those used for stock-market data. However, due to the different trading structures, tail behaviour, liquidity and investor composition of cryptocurrencies and stocks, this result cannot be directly applied to stocks.

Regime-based clustering will be applied to the other factors. Abdollahi and others construct a political-news index using natural language processing, quantify the volatility relationship between it and forty foreign asset market segments, and group the markets by sensitivity [12]. This study groups markets instead of dividing time into market regimes, but it can also show changes in cross-market structure through volatility connectedness and external information. Mudarisov et al. Use

DOI: 10.54254/2753-8818/2026.37205

clustered market regimes and FinBERT-based news representations to predict cross-sectoral regimes [13]. Based on the above results, text can help identify changes in the market and improve regime forecasting.

These studies explain the reasons for the classification differences. Assets are indices of stocks and shares, foreign exchange, Bitcoin, etc. Frequencies range from macro-economic monthly data to daily returns and high-frequency market microstructure data. Feature sets include marginal returns, rolling volatility, correlations, macroeconomic variables, option-implied densities and news. Some models need a predetermined state number; others infer the number of states or compare different specifications. Some ways assign each observation to one cluster, and others provide probabilities. Finally, the validation may be statistical separation, forecasting, portfolio results, temporal persistence or alignment with known crisis periods.

Therefore, the existing literature does not provide a universal "best" model. When research focuses on transparency and speed, K-means is still suitable; distributional distances enhance the recognition capability for changing return behaviour; mixture models can express ambiguous boundaries; HMMs can capture temporal persistence; jump models reduce some implausible switching; and rolling procedures adapt to non-stationarity. The Selection of a model should be based on the research question, data frequency, available features and desired interpretation, rather than on algorithmic complexity.

## 5. Limitations and prospects

The first is that there are no ground-truth labels. Market regimes are analytical constructs, not directly observable phenomena. Different studies may divide the same data into two or three different regimes and give different names to statistically similar states. State labels may also change due to random initialisation or rolling samples. Therefore, a high internal clustering index does not mean that the regime has a stable economic effect.

Second, the results of feature selection differ. Standardisation changes the relative scale of variables, window length affects the trade-off between responsiveness and noise, and highly correlated magnitude-related features may repeatedly encode the same dimension of risk. Extreme observations present another problem: if they are removed, the risk-generating events that the model needs to find will be excluded; otherwise, incorrect data may shift the centroid and covariance. Therefore, the preprocessing decisions and some sensitivity tests should be reported by the researchers instead of presenting a single pipeline as objective.

Third, financial data are non-stationary, and crisis samples are scarce. Models trained on a long period of stability may fail to reflect extreme-case behaviour, and models trained during a crisis may overfit that specific situation. Cross-market transfer is also unreliable. A Bitcoin volatility regime [11] may not be the same as a foreign-exchange regime because of differences in trading structure and the origin of risk for these two markets. Rolling HMM procedures can adapt to changes in the data [7], but retraining also has problems such as label alignment, computational cost and model selection.

Fourth, model assumptions can conflict with real market behavior. K-means favours compact clusters; Gaussian mixture models may not fit heavy-tailed returns well under restrictive component assumptions; and a first-order HMM assumes that the next regime depends only on the current state rather than the full history. More flexible methods can address the problems above, but they are more complex and less transparent. Some studies that build investment strategies based on detected states may be more susceptible to transaction costs, delayed signals, data snooping and look-ahead bias. Therefore, economic usefulness should be judged based on genuinely out-of-sample data.

DOI: 10.54254/2753-8818/2026.37205

Future research will develop a dynamic online update method to update the state after adding new data without arbitrarily redefining historical labels. Future work will propose probabilistic assignments to address the uncertainty near the transition. Multi-scale models can incorporate both short-term and long-term volatility windows, and multi-asset systems may include returns, correlations, liquidity, credit spreads, implied volatility and macroeconomic indicators. Textual information from political news and FinBERT applications [12, 13] can provide supplementary signals of changes in market conditions, but care must be taken to align them with the publication time to avoid future information leakage.

Strengthen the norm enforcement of evaluation. Research reports should present some key indicators of all regimes, such as state sizes, average durations, transition frequencies, preprocessing rules, initialisation procedures and sensitivity to alternative values of K. If possible, they should also reserve a later time period for out-of-sample testing and compare the economic interpretations of different regimes based on this, rather than only using the clustering index. The common reporting standards will not offer actual labels; rather, they may help us learn about and reproduce model differences.

Interpretability and robustness should also remain central. Researchers should describe the meaning of each state in terms of return, volatility, drawdown, correlation, and duration; compare results across seeds, windows, features, and periods; and test whether identified stress states remain recognizable across markets. Instead of claiming a final objective partition, future studies should treat regimes as uncertain and model-dependent summaries of evolving financial structure. This perspective would make regime detection more statistically transparent and more economically useful.

## 6. Conclusion

This review examined the use of volatility features and unsupervised learning for financial market regime detection. Financial prices can be transformed into returns, rolling volatility, drawdown, trading volume, implied volatility, and dynamic correlations, but the resulting regimes depend strongly on preprocessing, feature scaling, and window selection. K-means and hierarchical clustering provide relatively transparent groupings, Gaussian mixture models represent uncertain membership, spectral clustering can identify non-spherical structures, and hidden Markov or jump models incorporate temporal persistence. Recent studies increasingly use multidimensional distributions, probabilistic assignments, rolling estimation, high-frequency observations, option-implied information, and textual features. These developments broaden the information available for regime identification, but they do not eliminate the absence of universally accepted ground-truth labels. Model outputs remain sensitive to the number of states, initialization, sample period, feature design, and validation criteria. Reliable studies should therefore report preprocessing rules, state sizes, average durations, transition frequencies, and sensitivity analyses across seeds, windows, features, and markets. Out-of-sample evaluation is also necessary when detected regimes are used in risk management, portfolio allocation, or trading applications, because transaction costs, delayed signals, and look-ahead bias can otherwise exaggerate economic usefulness. Future research should develop online updating procedures that preserve label consistency, quantify uncertainty near regime transitions, and test whether stress states remain recognizable across asset classes and market periods. The main value of regime detection is not to assign an unquestionable label to every trading day. Its value lies in converting evolving financial behavior into structured and testable descriptions of market risk that support more transparent model evaluation and more robust financial decision-making.

DOI: 10.54254/2753-8818/2026.37205

## References


[1] Ang, A., & Timmermann, A. (2012). Regime changes and financial markets. *Annual Review of Financial Economics, 4*, 313–337. https://doi.org/10.1146/annurev-financial-110311-101808

[2] Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica, 57*(2), 357–384. https://doi.org/10.2307/1912559

[3] Prakash, A., James, N., Menzies, M., & Francis, G. (2021). Structural clustering of volatility regimes for dynamic trading strategies. *Applied Mathematical Finance, 28*(3), 236–274. https://doi.org/10.1080/1350486X.2021.2007146

[4] Luan, Q., & Hamp, J. (2025). Automated regime classification in multidimensional time series data using sliced Wasserstein k-means clustering. *Data Science in Finance and Economics, 5*(3), 387–418. https://doi.org/10.3934/DSFE.2025016

[5] Bucci, A., & Ciciretti, V. (2022). Market regime detection via realized covariances. *Economic Modelling, 111*, Article 105832. https://doi.org/10.1016/j.econmod.2022.105832

[6] Aydınhan, A. O., Kolm, P. N., Mulvey, J. M., & Shu, Y. (2024). Identifying patterns in financial markets: Extending the statistical jump model for regime identification. *Annals of Operations Research*. https://doi.org/10.1007/s10479-024-06035-z

[7] Pagliaro, A. (2026). Regime-aware LightGBM for stock market forecasting: A validated walk-forward framework with statistical rigor and explainable AI analysis. *Electronics, 15*(6), Article 1334. https://doi.org/10.3390/electronics15061334

[8] Oliveira, D. C., Sandfelder, D., Fujita, A., Dong, X., & Cucuringu, M. (2026). Tactical asset allocation with macroeconomic regime detection. *Quantitative Finance, 1*–27. https://doi.org/10.1080/14697688.2026.2659195

[9] Kim, M. J., & Kwon, D. (2023). Dynamic asset allocation strategy: An economic regime approach. *Journal of Asset Management, 24*(2), 136–147. https://doi.org/10.1057/s41260-022-00296-8

[10] Koukorinis, A., Peters, G. W., & Germano, G. (2025). Generative-discriminative machine learning models for high-frequency financial regime classification. *Methodology and Computing in Applied Probability, 27*(2), Article 36. https://doi.org/10.1007/s11009-025-10148-8

[11] Ardia, D., Bluteau, K., & Rüede, M. (2019). Regime changes in Bitcoin GARCH volatility dynamics. *Finance Research Letters, 29*, 266–271. https://doi.org/10.1016/j.frl.2018.08.009

[12] Abdollahi, H., Junttila, J.-P., & Lehkonen, H. (2024). Clustering asset markets based on volatility connectedness to political news. *Journal of International Financial Markets, Institutions and Money, 93*, Article 102004. https://doi.org/10.1016/j.intfin.2024.102004

[13] Mudarisov, T., State, R. V., Kräussl, Z., Yakubov, A., & Petrova, T. (2024). Cross-sector market regime forecasting with LLM-augmented news analysis. In *Proceedings of the 5th ACM International Conference on AI in Finance* (pp. 461–468). ACM. https://doi.org/10.1145/3677052.3698642