

# A Hybrid Learning Approach to Detecting Regime Switches in Financial Markets

Akiyamen, Tang, and Hussien (2021) - Western University

A Referee Report by Ricardo Semião e Castro

2025-04-30

## Overview

The paper proposes a novel methodology for identifying regime switches in financial markets by combining supervised and unsupervised learning techniques. The authors utilize a PCA-based dimensionality reduction approach followed by clustering to detect regimes in the US economy. Then, they create classification models to predict these regimes. Finally, they show the usefulness of the regime identification by backtesting portfolio strategies based on the predicted regimes. The data is daily percentage changes of macroeconomic variables from the Federal Reserve Economic Data (FRED) database, on the window from 1994 to 2020.

The authors claim that the literature on regime detection is more focused on traditional statistical methods, to which they want to contribute with their unsupervised approach, where the number of regimes is chosen passively.

The paper is short, and does not pose itself to explore all the complex modeling decisions. With this in mind, many comments below should be taken with this in mind.

## Major Comments

### Hybrid Approach

While indeed less work has been done with unsupervised regime detection, it is not novel. The hybrid approach of training a classification model afterwards is more original. However, its benefit of “a classification agnostic on how the regimes were defined” is not much useful in a practical sense, while it adds an additional prediction error in comparison to using the clustering predictor.

If someone were to use the paper’s methodology, it is guaranteed that they’d have the clustering model on hand to use. If not, then they are not using the paper’s methodology.

A possible benefit of the hybrid approach is to “purposely” add a prediction that will not match 100% the clustering prediction, hoping that this would be beneficial, but the authors do not show any evidence of this.

## Lack of Goal

The paper lies in a gray area between a methodological and an applied one. With this, it lacked a clear goal. What was the objective of the regime identification and prediction:

1. Was it to find agnostic regimes, that showed overall differences in the macroeconomy?
2. Or to find specialized regimes that would maximize the returns of a portfolio?

Each goal implies very different modeling decisions on variable selection and hyperparametrization. For instance, with 1., one would need to select a balanced set of variables among the different sectors of the economy, to avoid building regimes that relate only to say, government securities.

With 2., one would have to consider the strategies' needs: do they need information on volatility or return levels regimes? Do they need to correlate with specific assets? And then, one could choose the clustering and classification hyperparameters to maximize the returns of the portfolio.

## Reproducibility

As said, the paper is very concise, in a way that its reproducibility is compromised. The authors do not provide the exact variables used, just the general topics that they fall into.

The hyperparameters of the clustering and classification models, while they have known defaults, are not discussed. For example, the distance of the K-Means, the weak learner of AdaBoost, etc.

Additionally, the 10-fold cross-validation was not fully described, was a blocking or traditional split used?

## Data Analysis and Transformation

Beyond interpreting the clustering, the authors do not provide any analysis of the data. Given the methodological nature of the paper, it is less important to report graphical analysis, but it is important to discuss variable transformation.

While a common transformation with financial variables, they do not explain why only considered the percent change variation. The variables might have important data on their levels, and some are stationary in them, another not covered topic.

## Minor Comments

1. The regime coloring in Figure 2. is using a continuous scale  $[1, 2]$ , where a discrete scale  $\{1, 2\}$  would be best.
2. The backtest results' graphs show too little information. They should at least color the line by regime. But also, could show the other variables, and maybe even the variables ordered by returns. Without this, it is very hard to grasp how the regime prediction contributed to the accumulated returns, which was one of the main goals.
3. Interpreting the regimes and discussing strategies that would work based on them was a very nice addition to the paper. Given the small space, it is ok to only discuss a few strategies, but for a proper test-drive of the method, a more systematic approach would be expected.
4. K-fold cross-validation and checking correlation between performance of the classifier and the portfolio built on it were good ideas. Without this notion of "understanding which metric is more relevant", testing all the classifiers seems like too many tests. Unfortunately, they were poorly executed:

- The performance metrics shown are the in-sample ones, which are much less informative than the out-of-sample ones. Some results even show perfect accuracy. This was a very poor practice.
- The chosen set of metrics was not discussed. Given that the goal was to check which correlate best, a larger one would've been expected.
- The actual correlation checks are presented in a simplistic way, with no table and not much discussion.