https://doi.org/10.53656/igc-2025.13

PREDICTION OF MARKET REGIME CHANGES USING MACHINE LEARNING MODELS

Yanka Aleksandrova
University of Economics, Varna (Bulgaria)
Svetlin Minev
University of Economics, Varna (Bulgaria)

Abstract. The purpose of this research is to build models for financial market regime prediction. The study applies supervised machine learning to forecast financial market regime switches using various indicators based on S&P 500 index. Results demonstrate that Random Forest and XGBoost achieved over 83% accuracy for 5-day ahead forecast, while LSTM scored 78%. Stacked Ensemble of XGBoost and Random Forest combines the strength of both single models and proved to be the best model with 84% accuracy, ROC AUC of 0.944 and excellent capabilities to detect all market regimes. Overall results show that machine learning enhance the forecast of regime changes and can be integrated in efficient algorithmic trading strategies.

Keywords: market regimes forecast; machine learning; algorithmic trading
JEL: G17, G11, C53

Introduction

Financial markets are known to be cyclical, oscillating between different regimes of expansion, contraction, and indecision. Each market regime is characterized by distinct return distributions, volatility patterns, durations, and cross-asset correlations. Accurate prediction of regime transitions is critical for the design of trading strategies, portfolio building and risk management (Panayotova, 2025). Failure to detect bull trend overextensions could lead to large drawdowns, while missing the end of the bear market could result in decreased profitability. Adding market regime detection to the money management strategy will enhance the position sizing allowing long only strategies to trade larger volume during bull markets, lower volumes during extended bull or sideways markets and stay on the sidelines during bear markets. To be effective in real market conditions, such methods must remain statistically robust while providing enough transparency to ensure consistent implementation and risk control.

Literature Review

Predicting shifts in stock returns, volatility, cross-asset correlations, or market regimes has long been central to portfolio construction, risk management, and trade timing (Petkov, 2025). Classical econometric models (e.g., regime-switching variants) have been increasingly improved by machine-learning (ML) methods that use non-linear relationships, interactions, and higher-dimensional signals. Five broad families of ML algorithms are used in the analysis and prediction of market regime changes: supervised classifiers, unsupervised clustering, regime-switching statistical models, deep learning neural networks, and probabilistic (Bayesian) models, each addressing the challenges of non-stationarity and ambiguous regime definitions in different ways.

Supervised classification tends to dominate in real applications where market state classifications are available. Tree ensembles (Random Forest, gradient boosting such as XGBoost) are widely used because they handle feature interactions, class imbalance, and heterogeneous inputs with strong out-of-sample stability. Support Vector Machines appear in binary bull–bear tasks or as components within hybrids. Several studies report that market regime detection by using such models can improve drawdown control and Sharpe ratios in, though reported accuracy and economic gains vary with the labelling scheme and evaluation design (Ndoutoumou et al., 2025; Shu et al., 2024).

By contrast, unsupervised clustering methods (e.g., PCA with K-means or hierarchical clustering) infer regimes directly from the structure of the data and only afterward attempt to predict the resulting cluster assignments (Mc Greevy et al., 2024). This approach reduces reliance on expert-defined labels but is prone to instability of clusters over time. Many studies treat market regimes as hidden states using Hidden Markov Models or related Markov-switching frameworks (Ndoutoumou et al., 2025). These methods describe regime changes through transition probabilities and intuitive dynamics, and they often act as a backbone that machine learning extends with richer predictors or feature selection.

More recently, deep learning methods have expanded the toolset used for analysing and predicting regime changes. LSTMs are applied to capture temporal dependencies, CNNs to detect local patterns in rolling windows, and Variational Autoencoders or Transformer-based meta-learners aim to adapt quickly when distributions shift. Studies report improvements in classification accuracy and a better grasp of extreme moves during regime changes. At the same time, performance remains sensitive to non-stationarity and labelling choices, and deep networks bring practical challenges in implementation and interpretability. Complementing these are probabilistic/Bayesian methods (variational inference, Gaussian mixtures, Bayesian model averaging) that deliver calibrated uncertainty and cluster-specific inferences, aiding risk budgeting and governance (Jiang, 2021; Zhou, 2025).

All these approaches use inputs such as market variables like returns, volatility (realized or implied), breadth, and liquidity, or economic indicators like yields, inflation, and growth, and even alternative data such as news sentiment, options surfaces, or crypto on-chain metrics. Regimes are labelled either with expert rules—moving averages, drawdown thresholds, volatility bands—or inferred statistically through clustering or structural break methods. A consistent lesson in the literature is that the choice of labels strongly shapes the results: different definitions can alter class balance and regime timing, which in turn affects not only metrics like accuracy, F1, and AUC, but also the economic value observed in trading simulations (Liu & Li, 2023; Huang, Z. et al., 2023; Bucci & Ciciretti, 2022).

Evaluation methods have changed toward rolling and walk-forward splits, along with cost-sensitive metrics that recognize the higher risk of missing a bear market compared with a false bull signal. Tests of market regime shifts usually include transaction costs, because frequent regime changes can decrease the nominal profits. For practical application of the concepts, transparency and auditability are critical (Chae et al., 2024). Most common challenges are overfitting to historical data, data snooping, survivorship bias, and fragile performance when regimes shift over time (Salehpour & Samadzamini, 2023). Studies that combine clearly defined labelling rules with robust tree ensembles and time-aware evaluation could reduce label ambiguity, handle class imbalance, and keep the link to economic outcomes clear (Biolcheva, 2025). These designs help balance classification performance, including for rare bear markets, while still being practical for deployment in trading and risk systems, addressing issues of non-stationarity, costs, and governance from the outset (Noor & Fatima, 2025).

Methodology

Before applying machine learning models, rules to identify market regimes are defined and applied. Rule-based market regime detection is based on clear, well-understood conditions, including volatility and moving averages thresholds, as well as drawdown levels, which contribute to easier interpretation of the detected regimes (Kole & van Dijk, 2017; Zegadło, 2022). Such methods can be easily historically aligned and backtested across various economic cycles and financial events and periods, enabling analysts and researchers to evaluate robustness during crisis, expansions and corrections (Huang et al., 2023). As a benchmark S&P 500 index is chosen as it provides an added layer of relevance. This index is a wide and diversified measure that captures the global economic and financial markets dynamics (Wang, 2021; Wysocki & Slepaczuk, 2024).

To ensure interpretable and historically consistent classification of market regimes a set of technical indicators of volatility, moving averages, and RSI are used. Medium- and long-term trends are captured in the moving averages (50 day moving average (MA50) and 200 day moving average (MA200); their cross is

widely used as an indicator of either bullish (MA50 > MA200) or bearish (MA50 < MA200) momentum. Relative Strength Index (RSI) over a 14-day window measures short-term market momentum and investor sentiment, where values above 51 mean strong performance beyond normal and values below 51 mean weak performance. For measuring the short-term market risk, a 20-day Volatility (Volatility20) is applied. Volatility is calculated as the standard deviation of 20 moving average of the Close price. Values below 0.015 are indicative of stable growth and values above 0.045 are indicative of turbulent and crisis-driven performance. The thresholds have been calibrated by experiment and checked against established financial market regimes in historical times, including the 2008 Financial crisis, the Eurozone debt crisis and the COVID-19 shock. Such alignment allows the rules to capture theoretical signals but also can be aligned with actual historical periods and empirical evidence in S&P 500, making the framework more robust, transparent, and practical to use.

The rules for defining the market regimes are as follows:

– BULL condition: MA50>MA200 AND RSI>51 AND Volatitlity20<0.015

– BEAR condition: MA50<MA200 AND RSI<51 AND Volatility20>0.045

– SIDEWAYS condition: IF NOT BEAR OR BULL

A visual representation of the S&P 500 index from 2000 to 2024, with each point of time color-coded according to the identified market regime is shown on Figure 1. Vertical dashed lines highlight major global events such as the 2008 Financial crisis, COVID-19 pandemic crash, the post-COVID rise, the 2022 Bear market, and the subsequent AI-driven rise. The chart demonstrates how rule-based regimes align with historically significant market changes and can provide interpretable insight into market regimes across longer economic periods.

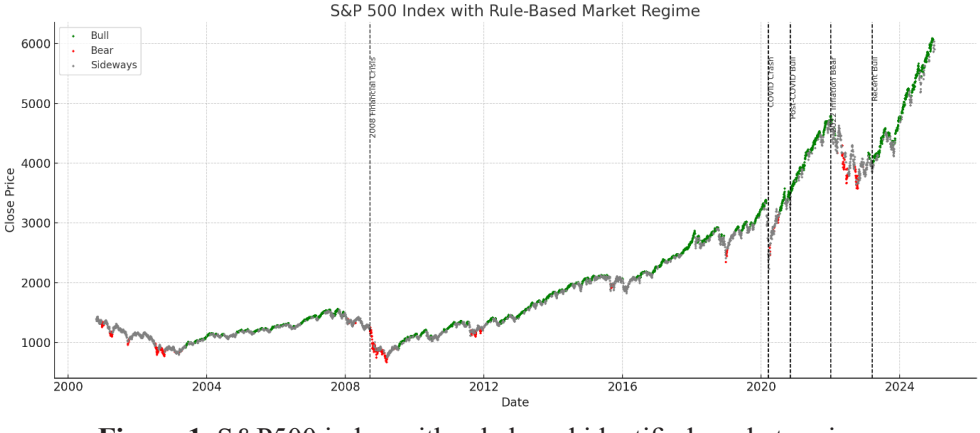


Figure 1. S&P500 index with rule based identified market regimes

A full outline of our proposed methodology for implementing machine learning models is given in Figure 2. The first step of the methodology is "Feature engineering", on which technical indicators are added to the dataset, such as:

– Bollinger Bands (BB) – this group of indicators is volatility-based and can be used to measure overbought and oversold levels. The Middle Band is 20 days Moving Average of S&P 500 price. We considered Upper BB, equal to Middle Band + 2*Volatility20, and Lower BB, calculated as Middle Band – 2*Volatility20. Prices below to the lower Bollinger Band suggest oversold assets, while prices above the Upper Band point at overbought ones.

– MA50 Slope – represents the rate of change of the 50 day Moving Average. This indicator measures short and midterm market trend strength and direction. A positive slope indicates upward momentum, while negative one reveals downward momentum. MA50_Slope is calculated as follows (1):

$$MA50_Slope = (MA50_t - MA50_{t-1})/MA50_{t-1} \quad (1)$$

– Daily Return – the percentage change in price from one day to the next one. This indicator captures short-term fluctuations and volatility and is calculated as follows (2):

$$r_t = P_t / P_{t-1} - 1 \quad (2)$$

– Volume Change –the relative change in traded volume compared to the previous day. Useful for measuring market activity and trading intensity.

– Price_to_MA50 – the relative distance of the price from its 50-day Moving Average, calculated as follows (3):

$$(Close_t - MA50_t) / MA50_t \quad (3)$$

– Price_to_MA200 – like the previous indicator, but for measuring long-term trend position. When the price is above 200-day Moving Average, the market can be identified as bullish.

– Relative Daily Range (High-Low Spread) – the difference between the highest and lowest price of the day, divided by the closing price. A wide range indicates high uncertainty and fluctuations within the day.

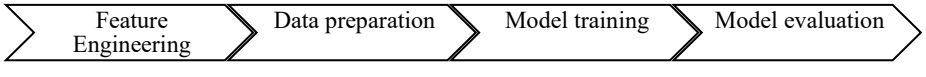


Figure 2. Methodology steps

On "Data preparation" step label encoding is performed for the target variable, while numerical features are standardized using Standard Scaler. As evidenced by

the chart on Figure 1, there are too little data points, identified as Bear markets. The class distribution of the dataset reveals a strongly imbalanced dataset, with 45.4% of cases belonging to the Bull class, 48.4% – to Sideways and only 6.1% of cases belonging to Bear class. This requires the implementation of appropriate techniques for balancing the training data in order to achieve models that can correctly predict all market regimes. The technique applied is Synthetic Minority Oversampling Technique (SMOTE) to balance the dataset prior the training, as this approach proved to be one of the best when dealing with strongly imbalanced datasets (Xu et al., 2019; Zhai et al., 2022; Pratheebha et al., 2024).

During the "Model training" several machine learning models are trained to predict market regimes 5 days ahead in order to transform from historical regime detection to taking action before the actual market regime shifts. A 5-day horizon is sufficient to use relevant signals, and not too long as to be providing uncertainty at the end of the horizon. The purpose of the models is to predict the future market regime, based on the technical indicators at the current point of time. Since this is a classification task, the models are trained using classification algorithms like Random Forest, XGBoost, LSTM (Long Short-Term Memory) neural network and Stacked Ensemble with best single models. During the training hyperparameter optimization is performed to identify the optimal values for model parameters.

After models are trained, they are evaluated and compared using metrics like Area Under Curve (AUC), Sensitivity, Accuracy, F1 score, etc. The comparative analysis of the models during the "Model evaluation" step assesses the predictive power of the trained models and is used to choose the best model.

Experimental results

According to the described methodology, all the models are trained on balanced dataset with hyperparameter optimization. To ensure equality, trained models are tested and evaluated with the same time-aware test dataset. Results from models' evaluation are presented in Table 1.

Table 1. Model evaluation results

Model	Accuracy	Macro F1	Macro Recall	ROC AUC
Random Forest	0.833	0.787	0.764	0.920
XGBoost	0.833	0.812	0.794	0.926
LSTM	0.775	0.756	0.754	0.512
Stacked Ensemble (RF+XGB)	0.844	0.820	0.837	0.942

The comparison of the various shows that there are significant differences in their performance in various metrics. Random Forest model has an accuracy of 0.833 and macro F1 score of 0.787. Although it is doing well in classifying the dominant classes (Bull and Sideways), it is doing poorly in the minority class

(Bear regime) where the recall is still relatively low at 0.597. Despite the ROC AUC of Random Forest is relatively high (0.920), evaluation results highlight that there is a weakness in Random Forest in being able to detect less common market regimes.

XGBoost is equally as accurate as Random Forest, but it surpasses the latter at macro F1 score with 0.812. This is because XGBoost is more effective at balancing performance between classes and scores higher on the recall of the minority regime (Bear). Its ROC AUC of 0.926 confirms that it distinguishes market regimes better than Random Forest across different thresholds. It is the most effective single model as it can deal with class imbalance in a more efficient way and reflects more complex patterns in engineered features.

The LSTM on the other hand performs significantly worse. It obtains the lowest accuracy of 0.775 and macro F1 of 0.756 and its ROC AUC decreases to 0.512, slightly higher than random guessing. This underperformance indicates that, in this case, a sequence model is not the best model to predict market changes. The usage of technical indicators like moving averages, volatility, and RSI are likely to reduce the LSTM advantages in capturing sequential dependence. Another possible explanation for the poor ability of the LSTM model to generalize is related to the dataset size and class imbalance. Overall, LSTM proved to be weaker than tree-based models across all evaluation metrics.

The Stacked ensemble model combining optimized Random Forest and XGBoost shows modest improvements over individual models. The accuracy of the Stacked Ensemble increases to 0.844, with a macro F1 score of 0.820. This effect can be attributed to the averaging of the probabilities of the two base learners, which allows the ensemble to reduce the variance and stabilize the predictions.

The evaluation results in Table 1 show that the Stacked Ensemble of Random Forest and XGBoost outperforms the other models. This model achieves the strongest recall across classes, with a macro recall of 0.837, demonstrating its ability to correctly identify even the rare regime transitions more effectively. The ROC-AUC measure of 0.942, also confirms that this model is the most reliable and balanced approach among the other trained models.

Overall, the evaluation shows that in the group of single models, the tree-based models, particularly XGBoost, are better for forecasting market regimes when using engineered technical indicators. Random Forest is performing relatively well but is weaker in the minority class detection. LSTM is not performing well and reveals weak predictive power due to problems with detecting sequential dependencies in the data. Stacked ensemble of XGBoost and Random Forest offer significant benefits by combining the strengths of these single models. The stacked ensemble model is the best-performing model, providing not only the highest accuracy and F1-score but also the strongest class detection as measured by ROC-AUC.

Discussion

The experimental results demonstrate that supervised learning algorithms have economically significant proper predictive capabilities of short-horizon regime labels. Practical experiments with S&P 500 data found the use of tree-based algorithms to be significantly more effective than sequence LSTM model. XGBoost was equally accurate to Random Forest, but had larger macro-F1 and ROC-AUC, which means it is a better regime classifier. The combination of the two tree learners (XGBoost and Random Forest) provided the most robust and consistent results, and increased macro recall and ROC–AUC and was feasible to identify minority regimes of the Bear type in early stages.

It's worth mentioning that there is a significant difference between single model results on most classes (Bull, Sideways) and the rarer Bear results. Random Forest still had comparatively lower recall of the minority regime even after resampling, but this was reduced by XGBoost, and most notably by the stacked ensemble. This is an important advantage as false negatives on Bear states are usually more expensive than false negative on Bull and Sideway regimes. The higher macro recall (0.837) and ROC AUC (0.942) mean that predictions can be converted into a signal which is more reliable to reduce risks in the event of a downside condition.

Some limitations can be considered that suggest directions for future work. The first limitation is related to the used data balancing technique. Although SMOTE has been used to solve class imbalance, synthetic oversampling can obscure local dependence in time-series settings. This suggests that other approaches such as class-weighted losses, focal loss or time-consistent resampling could be worth considering. The second limitation relates to the use of a single, although broad index, like S&P 500. Future analysis should be focused on international equities, interest rates, credit and commodities which would be used to test the methodology strength. Another consideration is the prediction period. Although the standard five-day horizon is operationally reasonable and convenient, it's not always optimal. Hence, additional model training in multi-horizon settings should be carried out to explore the predictive power and robustness of the models.

Conclusion

This study shows that supervised learning can provide timely and interpretable signals of market regime changes when trained on engineered technical indicators. Tree-based models, particularly XGBoost, and Stacked Ensemble outperform a sequence model in both overall accuracy and classification of the minority-class, with the ensemble achieving the highest macro recall and ROC – AUC on a time-aware test set. Future work should broaden asset coverage, explore alternative data balancing techniques, more market regime states and time horizons, and connect classification gains to trading strategies under realistic limitations. These future research directions would allow to transform from correct market regime labelling to robust, deployable investment and risk-management trading strategies.

Acknowledgement

This research is funded by the scientific project PNI KS24-02-PMB-RD "Methodologies for Predictive Modelling in a Business Context".

REFERENCES

Biolcheva, P. (2025). Artificial intelligence as a risk treatment tool in the circular economy. *Industrial Growth Conference 2024*, 49 – 59. https://doi.org/10.53656/igc-2024.05.

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

Chae, J., Jang, B-G. & Kim, T., (2024). The effect of regime-switching transaction costs and cash dividends on liquidity premia. *International Review of Financial Analysis*, *93*, 103186, https://doi.org/10.1016/j.irfa.2024.103186.

Huang, Z., Li, N. Mei, W., & Gang, W. (2023). Algorithmic trading using combinational rule vector and deep reinforcement learning. *Applied Soft Computing*, *147*, 110802. https://doi.org/10.1016/j.asoc.2023.110802.

Jiang, W. (2021). Applications of deep learning in stock market prediction: recent progress. *Expert Systems with Applications*, *184*, 115537. https://doi.org/10.1016/j.eswa.2021.115537.

Kole, E. & van Dijk, D. (2017). How to identify and forecast bull and bear markets? *Journal of Applied Econometrics*, *32*(1), 120 – 139. https://doi.org/10.1002/jae.2511.

Liu, P., & Li, J. (2023). Segment regression model average with multiple threshold variables and multiple structural breaks. The *Canadian Journal of Statistics*, *52*(1), 131 – 161. https://doi.org/10.1002/cjs.11764.

Mc Greevy, J., Muguruza, A., Issa, Z., Salvi, C., Chan, J. & Zuric, Z. (2024). Detecting multivariate market regimes via clustering algorithms. *SSRN*. https://dx.doi.org/10.2139/ssrn.4758243.

Ndoutoumou, J., Yin, Z. & Cheng, X., (2025). HMM-based market regime detection with RL for portfolio management. *IEEE 11th International Conference on Intelligent Data and Security (IDS)*, 68 – 70. https://doi.org/10.1109/IDS66066.2025.00022.

Noor, K. & Fatima, U. (2025). Meta learning strategies for comparative and efficient adaptation to financial datasets. *IEEE Access*, 13. 24158 – 24170.

Panayotova, T. (2025). Preconditions for adequate approximation of the company value. *Industrial Growth Conference 2024,* 386 – 396. https://doi.org/10.53656/igc-2024.28.

Petkov, K. (2025). Logical framework method for conducting pre- and post-money valuations of equity crowdfunded companies leveraging AI. *Industrial Growth Conference 2024,* 207 – 219. https://doi.org/10.53656/igc-2024.14

Pratheebha, T., Megala, S. S., & Indumathi, V. (2024). An improved churn prediction model using relevance feature identification and personalized modeling technique with optimized retention strategies. *8th International Conference on*

Electronics, Communication and Aerospace Technology (ICECA), 1489 – 1496. https://doi.org/10.1109/ICECA63461.2024.10800772.

Salehpour, A., & Samadzamini, K. (2023). Machine learning applications in algorithmic trading: A comprehensive systematic review. *International Journal of Education and Management Engineering*, *13*(6). https://doi.org/10.5815/ijeme.2023.06.05.

Shu, Y., Yu, C., & Mulvey, J. M. (2024). Downside risk reduction using regime-switching signals: a statistical jump model approach. *Journal of Asset Management, 25,* 493 – 507.

Wang, R. (2021). *Discriminating modelling approaches for point in time economic scenario generation*. [Master's Thesis]. ETH Zürich. https://arxiv.org/pdf/2108.08818.

Wysocki, M., & Slepaczuk, R. (2024). Construction and hedging of equity index options portfolios. *SSRN*. https://dx.doi.org/10.2139/ssrn.4897319.

Xu, X., Chen, W., & Sun, Y. (2019). Over-sampling algorithm for imbalanced data classification. *Journal of Systems Engineering and Electronics, 30*(6), 1182 – 1191 https://www.jseepub.com/EN/10.21629/JSEE.2019.06.12.

Zegadło, P. (2022). Identifying bull and bear market regimes with a robust rule-based method. *Research in International Business and Finance*, *60*, 101603. https://doi.org/10.1016/j.ribaf.2021.101603.

Zhai, J., Qi, J. & Shen, C. (2022). Binary imbalanced data classification based on diversity oversampling by generative models. *Information Sciences*, 585, 313 – 343. https://doi.org/10.1016/j.ins.2021.11.058.

Zhou, Z. H. (2025). *Ensemble learning: Foundations and algorithms*. (2nd ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781003587774.

✉ **Assoc. Prof. Yanka Aleksandrova, PhD**
ORCID iD: 0000-0002-4165-3011
Department of Informatics
University of Economics – Varna
77, Knyaz Boris I Blvd.
9902 Varna, Bulgaria
E-mail: yalexandrova@ue-varna.bg

✉ **Assist. Svetlin Minev**
ORCID iD: 0009-0008-5592-547X
Department of Industrial Business and Logistics
University of Economics – Varna
77, Knyaz Boris I Blvd.
9902 Varna, Bulgaria
E-mail: svetlin.minev@ue-varna.bg