MARKET REGIMES AND PORTFOLIO ALLOCATION: EVIDENCE FROM THE ROMANIAN EQUITY MARKET USING HIDDEN MARKOV MODELS AND XGBOOST

Bogdan Ionut ANGHEL, PhD Candidate[*]
Georgi MARINOV, PhD[**]
Paul Cristian DONOIU, PhD Candidate[***]

Abstract

This study examines whether the Romanian equity market exhibits persistent risk regimes, whether these regimes can be forecast at short horizons, and whether such forecasts can improve portfolio allocation. The analysis uses daily BET index data over 2016–2025 and applies a two-step empirical framework. A Hidden Markov Model is first used to identify latent market states based on returns and realised volatility, followed by an XGBoost classifier that predicts regimes one day ahead using market, volatility, domestic macro-financial, and global indicators. The results reveal three distinct regimes, interpreted as calm, stress, and panic, each reflecting different levels of market risk. The forecasting model performs well in the calm and stress regimes, while predictions for the panic state are less precise but remain informative from a risk-management perspective. Most classification errors occur between neighbouring regimes rather than between extreme states. The economic evaluation shows that a regime-probability portfolio strategy with volatility targeting outperforms a passive Buy-and-Hold investment in the BET index, generating higher cumulative performance and lower drawdowns. Overall, the findings indicate that risk regimes in the Romanian equity market are persistent, short-term predictable, and economically useful for systematic portfolio allocation.

[*] *Department of International Business and Economics, Bucharest University of Economic Studies, Romania. ORCID: https://orcid.org/0000-0002-7043-2226. Corresponding author, e-mail: bogdan.anghel@rei.ase.ro*

[**] *University of Economics - Varna, Bulgaria. ORCID: https://orcid.org/0000-0002-2300-5635.*

[***] *Department of International Business and Economics, Bucharest University of Economic Studies, Romania. ORCID: https://orcid.org/0009-0003-0953-2759.*

Keywords: machine learning, BET Index, regime forecasting, volatility targeting

JEL Classification: C53, C58, G11, G15

DOI: https://doi.org/10.65672/fs.2026.2.4

1. Introduction

A central tension in financial economics lies between the premise of informational efficiency and the empirical observation that risk is time-varying. In its canonical form, the Efficient Market Hypothesis implies that systematic predictability in returns should be limited (Fama, 1970). Yet even in highly developed markets, return distributions exhibit volatility clustering, heavy tails, and episodic crises that differ qualitatively from tranquil phases. These features challenge the adequacy of models that treat the data-generating process as stable through time. In emerging equity markets, such departures from stability are often magnified. Lower liquidity, narrower investor bases, and stronger sensitivity to macroeconomic and policy shocks create conditions in which returns and risk may shift discontinuously across states rather than evolving smoothly (Bekaert & Harvey, 1997).

This paper studies regime-dependent risk dynamics in the Romanian equity market, proxied by the BET index, over 2016–2025. The sample contains multiple episodes associated with pronounced risk repricing, particularly the COVID-19 shock in 2020 and the tightening-driven stress of 2022. The realised volatility profile of BET indicates substantial time variation and clustering, with spikes that are difficult to reconcile with a single stable risk environment. Such dynamics motivate a regime perspective: if the market alternates between latent states—here conceptualised as calm, stress, and panic—then both inference and portfolio decisions may benefit from explicitly modelling the state.

Although regime-switching models provide a well-established statistical framework for capturing state dependence (Hamilton, 1989), and the portfolio-choice literature emphasises that optimal allocation can be conditional on regimes (Ang & Bekaert, 2002; Guidolin & Timmermann, 2007), evidence for smaller emerging European markets remains relatively limited. Moreover, two gaps persist in empirical practice. First, many regime studies prioritise statistical identification without demonstrating economic usefulness under realistic out-of-sample decision rules. Second, machine learning applications in finance often report predictive metrics without tying forecasts to transparent portfolio mechanisms and evaluating economic outcomes. This paper addresses these gaps by integrating regime extraction, regime forecasting, and implementable allocation within a single framework and judging success by both statistical diagnostics and economic performance.

The empirical strategy follows a deliberate separation of tasks. Latent regimes are first identified via a Hidden Markov Model (HMM) estimated on BET returns and realised volatility. Next-day regimes are then forecast using a multi-class gradient boosting classifier (XGBoost) trained on a conservative predictor set that combines volatility measures, lagged asset returns, and domestic macro-financial changes, with model complexity controlled through sequential feature selection and time-series cross-validation. Finally, predicted regime probabilities are translated into portfolio decisions through a regime-probability allocation rule that forms state-contingent return and covariance estimates and scales exposure via volatility targeting. This design reflects the view that forecasting the risk environment may be more stable and economically actionable than predicting return magnitudes, especially in markets where tail events drive welfare and long-run compounding.

The analysis is guided by five hypotheses. First, the Romanian equity market exhibits statistically distinct latent regimes. Second, these regimes display persistence rather than random day-to-day switching. Third, observable financial and macro-financial variables contain information about next-day regime transitions. Fourth, a nonlinear classifier can discriminate between regimes with economically meaningful out-of-sample accuracy, particularly by avoiding catastrophic errors that confuse panic with calm. Fifth, a regime-probability allocation strategy based on predicted states outperforms a passive Buy-and-Hold benchmark in cumulative and downside-risk terms.

2. Literature review

Research on financial returns has long documented systematic deviations from homoskedastic Gaussian benchmarks. Volatility clustering and heavy tails are among the most robust empirical regularities. The ARCH model introduced by Engle (1982) and its generalisation in the GARCH framework (Bollerslev, 1986) formalise conditional heteroskedasticity by allowing volatility to depend on past shocks. While these models capture smooth persistence in variance, they do not by themselves resolve a deeper issue: the mapping from information to risk may itself shift over time. When markets transition between environments governed by different constraints, funding conditions, or investor risk tolerance, a single stable conditional-variance law may be an incomplete representation. In such settings, volatility persistence can partly reflect the mixture of distinct regimes rather than within-regime dynamics.

This observation motivates regime-switching approaches. Hamilton (1989) provides a Markov-switching framework in which parameters depend on an unobserved state following a first-order Markov process. In finance, regimes are often interpreted as economically meaningful phases such as calm, stress, and crisis, potentially associated with changes in volatility, correlation structure, liquidity, and risk premia. Ang and Timmermann (2012) highlight that predictability is frequently conditional: relationships that appear weak or unstable unconditionally can be stronger within regimes, and models can fail when they ignore state dependence.

Regime considerations are particularly salient in emerging markets. Bekaert and Harvey (1997) show that emerging equity markets exhibit higher volatility and stronger sensitivity to global conditions, consistent with partial integration and capital flow dynamics that can amplify shocks. Such markets may exhibit nonlinear response patterns: similar news can induce modest adjustments in tranquil times but trigger outsized repricing when liquidity is thin or when investors are already risk-averse. This macro-financial amplification logic implies that regimes may correspond to different configurations of funding conditions and risk-bearing capacity, not merely different levels of recent volatility.

State dependence has direct implications for portfolio choice. Traditional mean–variance theory treats expected returns and covariances as stable inputs (Markowitz, 1952). If moments differ across states, unconditional optimisation can be misleading. Ang and Bekaert (2002) and Guidolin and Timmermann (2007) demonstrate that regime-switching representations can materially affect optimal allocation, especially when regimes are persistent and differ in tail risk. Importantly, economic value can arise even from coarse state awareness, because avoiding deep drawdowns increases geometric compounding—an effect that can dominate long-run outcomes.

A further strand concerns the role of nonlinear forecasting tools. Machine learning methods offer flexible function approximation capable of capturing interactions and nonlinearities (Hastie, Tibshirani and Friedman, 2009). In asset pricing and prediction, Gu, Kelly and Xiu (2020) show that machine learning can improve performance when signals are weak and dispersed across predictors. Yet the field is clear that financial data are noisy and that rigorous out-of-sample design is essential to avoid spurious patterns (Campbell, Lo and MacKinlay, 1997). These considerations motivate a conservative approach: model complexity should be constrained, and the value of forecasts should be evaluated by their economic consequences rather than in-sample fit.

This paper integrates this literature through a three-part argument. First, emerging markets plausibly exhibit discrete latent risk regimes due to macro-financial sensitivity and liquidity constraints. Second, regime-switching frameworks provide a coherent statistical mechanism for extracting such states. Third, nonlinear supervised learning may capture state-dependent signals embedded in market and macro-financial indicators, but must be evaluated out-of-sample and judged by economic performance within transparent allocation rules.

3. Methodology

The empirical analysis uses daily-frequency time series data sourced from Refinitiv for 2016–2025. The central market variable is the ROMANIA BET INDEX – PRICE INDEX, which proxies Romanian equity market performance. The dataset also contains daily financial series (e.g., global volatility indices and exchange rates where available), commodity prices, and a set of domestic macro-financial indicators. Because macro series are typically released at lower frequency, they are transformed into growth rates (percentage changes) and converted to daily panels through forward filling after resampling, consistent with the informational structure of lower-frequency releases.

The forecasting exercise is conducted under a strict chronological design. The training sample ends on 31 December 2022, and the out-of-sample test period begins on 1 January 2023. All time series are aligned to a common trading calendar to ensure temporal consistency across variables. Observations with missing values—arising primarily from rolling-window calculations or the asynchronous release of macroeconomic indicators—are excluded whenever they result in incomplete feature vectors. Asset returns are computed as continuously compounded (log) returns, following the standard approach in empirical financial research (Campbell, Lo and MacKinlay, 1997).

In addition to returns, several volatility and market-based indicators are constructed. These include 21-day and 63-day realised volatility, annualised using the square-root-of-time rule, as well as an exponentially weighted volatility measure. Additional variables capture different aspects of market dynamics, including absolute returns, rolling skewness and kurtosis calculated over a 21-day window, momentum over 21 and 63 trading days, and the difference between the 50-day and 200-day moving averages. A jump indicator is also included, defined as a dummy variable identifying unusually large price movements relative to recent volatility.

The empirical strategy proceeds in three steps: (i) infer latent regimes with an HMM; (ii) forecast next-day regimes with XGBoost; and (iii) translate predicted regime probabilities into a portfolio rule and evaluate economic outcomes relative to Buy-and-Hold. Separating regime identification from regime forecasting reduces the risk that the prediction model is merely fitting the object it aims to forecast and supports a clearer economic interpretation of the target.

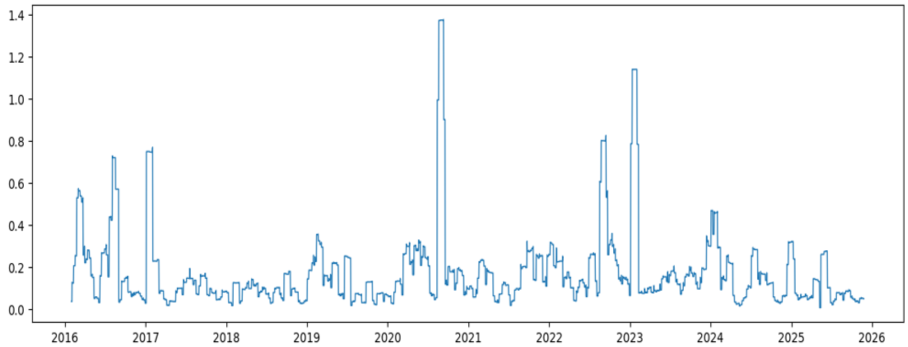
The Hidden Markov Model assumes that observed returns are generated by one of K latent states and that state transitions follow a first-order Markov process. Parameters are estimated by maximum likelihood using the Expectation - Maximisation algorithm (Baum et al., 1970; Rabiner, 1989). The implementation uses $K = 3$ states, estimated on the training sample, and then applies the fitted model to infer state assignments across the full panel.

The regime-extraction feature set includes BET log returns and 21-day realised volatility. Where present in the dataset, additional risk indicators are included (e.g., VIX, V2TX, MOVE, Romanian CDS, EUR/RON). Features are standardised using a scaler fit on the training sample and then applied to all dates to avoid look-ahead bias. The three states are mapped to economic labels (calm, stress, panic) according to the average realised volatility within each state: the lowest-volatility state is labelled calm, the intermediate state stress, and the highest-volatility state panic.

Figure 1 reports 21-day realised volatility (annualised) for BET. The series exhibits strong clustering and episodic spikes. The most pronounced surge is associated with the 2020 pandemic shock, while 2022 displays a sustained elevation in volatility consistent with a prolonged risk repricing episode. These dynamics are consistent with stylised facts documented by conditional heteroskedasticity models (Engle, 1982; Bollerslev, 1986), but they also suggest the plausibility of discrete shifts between qualitatively different risk environments. In this setting, modelling the market as switching across latent regimes is not merely a statistical convenience: it reflects the economic intuition that the constraints and risk-bearing capacity of investors can change abruptly under stress.

Figure 1

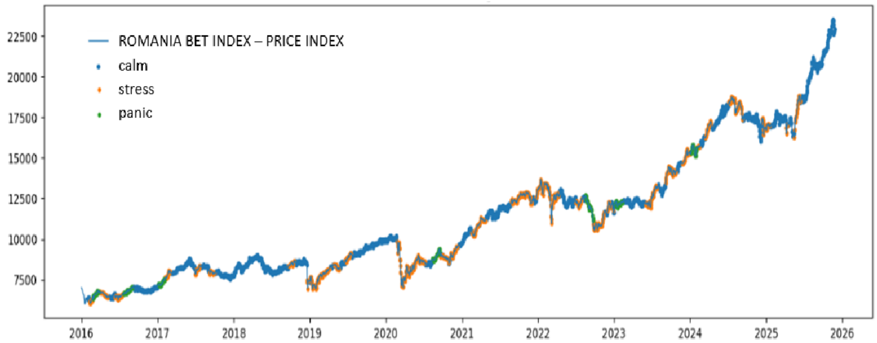
Realised volatility of BET (21-day, annualised)



Source: Author's contribution

The inferred market regimes display clear differences in the behaviour of the BET index. Periods classified as calm tend to coincide with sustained upward movements and relatively stable price dynamics. In contrast, stress regimes are more frequently observed around intermediate corrections and episodes of elevated uncertainty, where market movements become less stable but do not yet reach extreme levels. Panic regimes appear much less frequently and are typically concentrated in short intervals associated with abrupt market disruptions. As illustrated in Figure 2, the fact that panic observations occur in distinct clusters rather than being scattered randomly throughout the sample provides additional support for the interpretation that the Hidden Markov Model captures economically meaningful latent states rather than merely reflecting statistical noise.

Figure 2
HMM-inferred market regimes mapped onto the BET price index



Source: Author's contribution

The supervised learning task is next-day regime prediction. The regime label at date t is shifted forward to construct a target variable representing the regime at $t + 1$. Predictors are dated t, ensuring the model is trained under a genuine forecasting configuration. This framing is deliberate: a regime label can be interpreted as a compact representation of the conditional distribution of returns and therefore is directly relevant for risk management decisions.

First, market-based risk and dislocation measures capture contemporaneous stress accumulation: realised volatility (21 and 63 days), EWMA volatility, absolute returns, rolling skewness and kurtosis, momentum, moving-average dislocation, and a jump proxy. Second, lagged log returns of a set of tradable assets (up to 12 "price-like" series in the dataset) capture short-horizon cross-market dynamics and spillovers. Third, domestic macro-financial indicators are transformed into growth rates and forward-filled to daily frequency, capturing changes in financial conditions and real-economy momentum that can co-move with risk regimes in emerging markets.

Regime forecasting is implemented as a multi-class classification problem using XGBoost, trained with the multi-class log-loss objective (Friedman, 2001; Chen & Guestrin, 2016). The model produces calibrated regime probabilities $(p_{\text{calm}}, p_{\text{stress}}, p_{\text{panic}})$, which are used for ROC analysis and for portfolio construction.

Given the high-dimensional feature set and the noisy nature of financial data, overfitting risk is mitigated through a staged procedure consistent with best practice (Hastie, Tibshirani and Friedman, 2009).

First, feature-family selection retains only the most important feature within each family (e.g., among lags of the same return series), reducing redundancy and collinearity. Second, SHAP-based ranking is used to retain the top 15 features, improving interpretability and further restricting dimensionality. Third, hyperparameters are tuned using time-series cross-validation (TimeSeriesSplit) with RandomizedSearchCV, using a negative log-loss scoring criterion, reflecting an emphasis on probabilistic quality rather than only hard labels. The final configuration reported in your outputs exhibits conservative regularisation and shallow depth, aligning with a stability-oriented bias–variance trade-off.

Economic value is assessed by mapping predicted regime probabilities into an implementable allocation rule. Rather than using a discrete, regime-dependent equity weight, your implementation constructs state-conditional return and risk moments and then scales risk through volatility targeting.

For each regime $r \in \{\text{calm,stress,panic}\}$, the strategy estimates regime-conditional mean returns μ_r and covariance matrices Σ_r for the asset return vector using historical observations assigned to that regime. At each date t, predicted probabilities $p_t(r)$ are used to form probability-weighted moments:

$$\mu_t = \sum_r p_t\,(r)\,\mu_r, \Sigma_t = \sum_r p_t\,(r)\,\Sigma_r. \quad (1)$$

Portfolio weights are then formed from a mean–variance signal proportional to $\Sigma_t^{-1}\mu_t$, normalised by the L1 norm to control weight concentration. Finally, exposure is scaled to target a fixed annualised volatility (12%), using a rolling estimate of realised portfolio risk, which operationalises a risk-budgeting discipline related to volatility targeting. Performance is compared to Buy-and-Hold in the BET index. Equity curves, drawdowns, and calendar-year returns are computed by compounding returns. In the present implementation, daily returns are treated as log returns and annual returns are computed by exponentiating annual log-return sums.

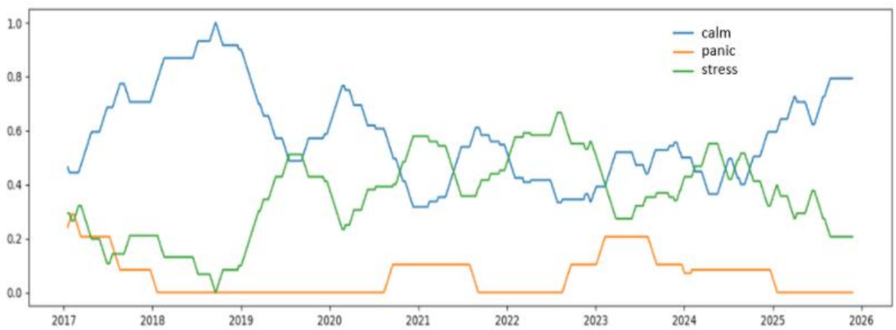
4. Results

This section presents the empirical results of the study and evaluates the extent to which market regimes in the Romanian equity market can be identified, forecasted, and exploited in a portfolio allocation framework. The analysis proceeds in several steps. First, the regime structure inferred from the Hidden Markov Model is examined in order to assess whether the estimated calm, stress, and panic states correspond to economically meaningful phases of market behaviour. Second, the predictive performance of the XGBoost classifier is evaluated using out-of-sample forecasts of next-day regimes. Particular attention is given to the model's ability to distinguish between different levels of market stress and to the economic implications of classification errors. Finally, the regime forecasts are incorporated into a probability-based portfolio allocation strategy, and the resulting investment performance is compared with a passive Buy-and-Hold benchmark in the BET index.

The evolution of regime frequencies over time provides additional insight into the dynamics of the market states. The evidence indicates that the distribution of regimes is far from constant across the sample period. Calm regimes tend to dominate during prolonged phases of market stability, whereas the share of stress regimes increases when uncertainty rises and market conditions become more fragile. Panic episodes remain relatively rare, but they reappear in clearly identifiable clusters rather than disappearing altogether. As illustrated in Figure 3, this variation in regime composition has important analytical implications. In particular, it suggests that unconditional return statistics effectively combine observations drawn from fundamentally different market environments. Consequently, analysing the data within a regime-dependent framework is likely to provide a more informative representation of market dynamics than treating the full sample as a single homogeneous process.

Figure 3

Rolling one-year regime frequency

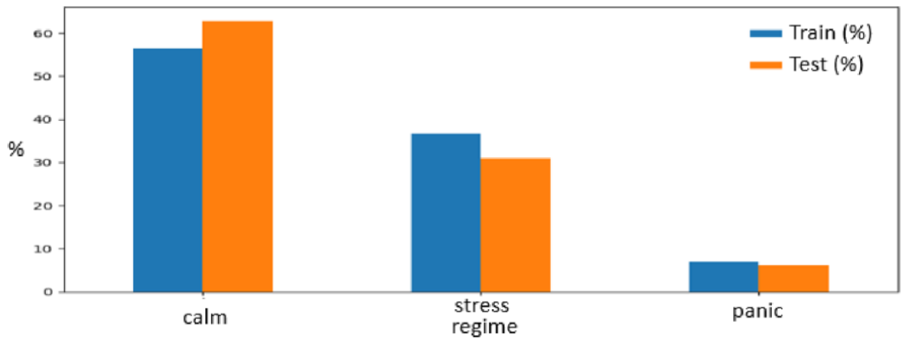


Source: Author's contribution

A comparison of regime frequencies between the training and test samples provides an additional perspective on the stability of the regime structure. Although the panic regime remains the least frequent state in both periods, its relative share is broadly comparable across the two subsamples. As shown in Figure 4, this similarity is particularly important for the interpretation of the forecasting results. It suggests that the out-of-sample evaluation period is not characterised by unusually benign market conditions. Instead, the test sample still contains a number of meaningful tail-risk episodes, which strengthens confidence in the robustness of both the regime prediction exercise and the subsequent portfolio performance analysis.

Figure 4

Regime distribution in training vs test subsamples



Source: Author's contribution

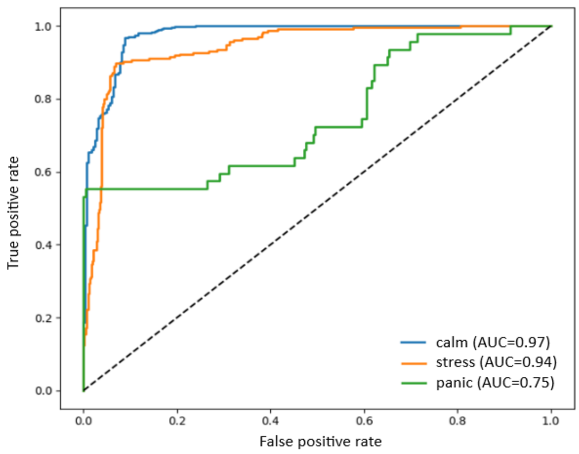
The second stage of the empirical analysis focuses on forecasting market regimes using an XGBoost classifier. The model is trained to predict the regime one day ahead, relying on the set of predictors selected through the SHAP-based feature selection procedure and hyperparameters determined via time-series cross-validation. The objective is to evaluate whether observable market and macro-financial indicators contain sufficient information to anticipate transitions between the calm, stress, and panic regimes previously identified by the Hidden Markov Model.

Out-of-sample results for the 2023+ test period indicate strong predictive discrimination between the calm and stress states. Predictive performance for the panic regime is weaker, though still informative, which is economically unsurprising given the rarity of extreme market events. Episodes classified as panic are typically associated with abrupt shocks that are only partially reflected in observable predictors. Importantly, the pattern of classification errors remains economically meaningful. Most misclassifications occur between neighbouring regimes, rather than between the extreme states of calm and panic. In practical terms, this implies that the model rarely interprets periods of severe market stress as tranquil conditions. From a portfolio management perspective, such an error structure is considerably less problematic than indiscriminate misclassification across the entire spectrum of market risk.

The predictive performance of the classifier is further evaluated using receiver operating characteristic (ROC) curves, reported in Figure 5. The results confirm strong discriminatory power for the calm and stress regimes. The calm state achieves an area under the curve (AUC) of 0.97, indicating near-perfect separation from the other regimes. Similarly, the stress regime records an AUC of 0.94, suggesting that the model effectively captures conditions associated with elevated, though not extreme, market volatility. Predictive accuracy for the panic regime is more moderate, with an AUC of 0.75. While lower than the other regimes, this level of performance still indicates that the classifier retains meaningful predictive ability even for the most extreme market state.

Figure 5

ROC Curves Test Final Model



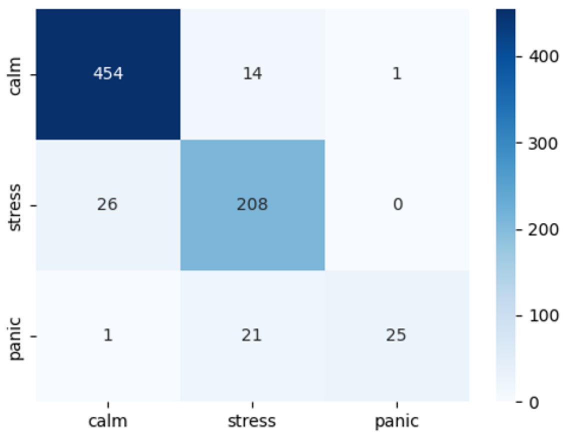
Source: Author's contribution

Additional insight into the model's performance can be obtained from the confusion matrix presented in Figure 6. The classifier performs very well in identifying calm market conditions, correctly classifying 454 observations, with only a few cases misclassified as stress (14) or panic (1). The stress regime is also identified with good accuracy, with 208 correctly classified observations, although some stress periods are predicted as calm (26). This likely reflects the fact that transitions between calm and moderately volatile market conditions tend to occur gradually.

The panic regime is more difficult to detect, with 25 observations correctly classified and several predicted as stress (21). However, the model almost never confuses panic with calm, with only one observation misclassified in this way. From a practical perspective, this pattern of errors is desirable because it reduces the risk of interpreting severe market stress as a tranquil market environment.

Figure 6

Confusion Matrix Test Final Model



Source: Author's contribution

The final model specification is obtained through a hyperparameter optimisation procedure applied after the SHAP-based feature selection stage. The search is conducted using time-series cross-validation combined with a RandomizedSearch approach. The selected hyperparameters are reported in Table 2. Overall, the resulting configuration relies on a relatively large number of trees with shallow depth and regularisation terms, which helps control model complexity and reduce the risk of overfitting in a financial dataset characterised by noisy signals.

Table 2

Final XGBoost Hyperparameter Configuration

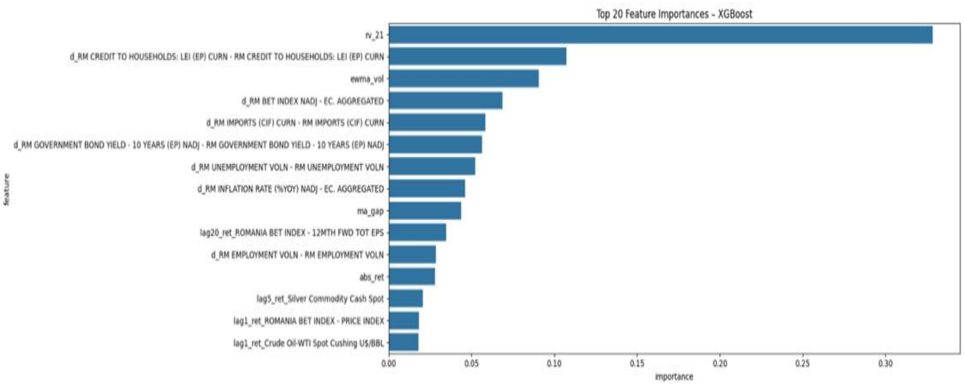
Hyperparameter	Value
n_estimators	800
max_depth	3
learning_rate	0.01
subsample	0.9
colsample_bytree	0.7
min_child_weight	1
gamma	1.0
reg_alpha (L1)	0.001
reg_lambda (L2)	1.0

Source: Author's contribution

Additional insight into the model's predictive performance is provided by the feature importance results shown in Figure 5. Realised volatility appears to be the most important predictor, consistent with the regime framework, since the different states mainly reflect changes in market risk. At the same time, the model also incorporates information from several domestic macro-financial variables, including changes in household credit and sovereign bond yields. This suggests that regime transitions are not driven only by recent market volatility, but may also be related to broader changes in financial conditions.

Figure 5

XGBoost feature importances for next-day regime prediction

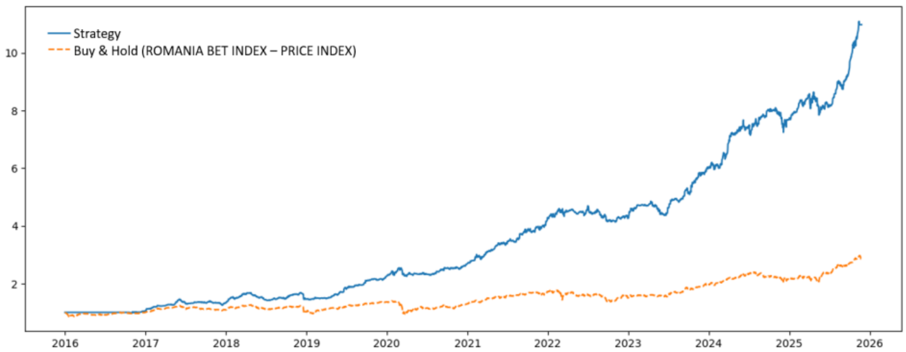


Source: Author's contribution

The final question is whether forecasting regimes lead to economically meaningful improvements in portfolio performance. Figure 6 addresses this directly by comparing the cumulative equity curve of the regime-probability strategy with that of a passive Buy-and-Hold position in the BET index. The difference is substantial. The strategy reaches a markedly higher terminal wealth level, and the gap does not emerge gradually in a trivial way; rather, it tends to widen after adverse market episodes. This pattern is consistent with the idea that limiting exposure during severe drawdowns materially improves long-run compounding, even if returns in normal periods are not dramatically different.

Figure 6

Cumulative equity curves: regime-probability strategy vs Buy-and-Hold

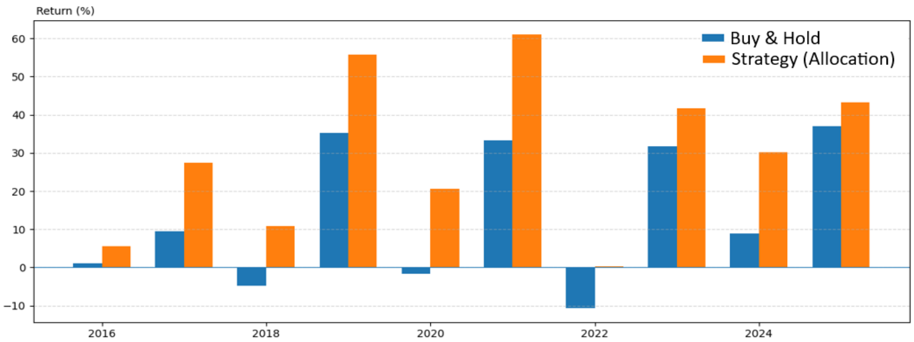


Source: Author's contribution

Annual returns are reported in Figure 7, which shows that the strategy outperforms the Buy-and-Hold benchmark in every calendar year between 2016 and 2025. This includes years when the market performs poorly, such as 2018, 2020, and especially 2022. These results suggest that the strategy is not simply defensive. Higher performance does not arise from persistently low exposure to risky assets. Instead, the strategy continues to participate in favourable market conditions while reducing exposure during periods of elevated risk. Such behaviour is consistent with what would be expected from a genuinely state-dependent allocation approach.

Figure 7

Annual returns: strategy vs Buy-and-Hold



Source: Author's contribution

5. Conclusion

The empirical analysis provides evidence that the Romanian equity market exhibits a regime structure that is both statistically identifiable and economically meaningful. Estimation of the three-state Hidden Markov Model reveals distinct calm, stress, and panic regimes, which correspond closely to observable shifts in realised volatility and coincide with well-defined periods of market turbulence. This alignment suggests that the inferred states capture genuine variation in the underlying risk environment rather than purely statistical segmentation of the return series.

The subsequent forecasting exercise evaluates whether transitions between these regimes can be anticipated using observable market and macro-financial information. The results indicate that the XGBoost classifier achieves a meaningful degree of out-of-sample predictive accuracy, particularly in distinguishing between the calm and stress states. Predictive performance for the panic regime is naturally weaker, reflecting its relatively low frequency and the fact that extreme market disturbances are often associated with abrupt exogenous shocks. Nevertheless, the pattern of classification errors remains economically informative. Misclassifications occur predominantly between adjacent regimes, implying that periods of severe market stress are rarely interpreted as tranquil conditions. From a portfolio management perspective, such an error structure is considerably less problematic than indiscriminate misclassification across the full range of risk states.

An examination of feature importance provides further insight into the drivers of regime forecasts. Realised volatility emerges as the dominant predictor, which is consistent with the conceptual foundation of the regime framework, where states are primarily differentiated by their risk characteristics. At the same time, several domestic macro-financial variables contribute non-negligibly to predictive performance. In particular, indicators related to household credit dynamics and sovereign yield movements appear among the most influential variables. This pattern suggests that regime transitions are not merely mechanical reflections of recent volatility, but are also associated with broader changes in financial conditions that influence market resilience and the propagation of stress.

The economic relevance of the forecasting framework becomes evident when regime predictions are incorporated into a portfolio allocation strategy. The implemented approach constructs regime-probability-weighted estimates of expected returns and covariance matrices, while overall exposure is controlled through a volatility-targeting mechanism. Over the full sample, this strategy generates a substantially stronger investment profile than a passive Buy-and-Hold position in the BET index. The cumulative performance indicates a markedly higher terminal wealth level, accompanied by materially smaller drawdowns. In addition, the strategy outperforms the benchmark in calendar-year returns throughout the 2016–2025 period, including years characterised by negative market performance.

While these results are encouraging, several considerations remain relevant for interpretation. The current implementation does not explicitly incorporate transaction costs or liquidity constraints, both of which may affect realised investment performance, particularly in emerging markets. Future work should therefore account for such frictions and evaluate the sensitivity of the results to realistic cost assumptions. In addition, further robustness analysis would be desirable, including alternative regime specifications, different predictive models, and rolling out-of-sample estimation windows.

Taken together, the findings suggest that the risk dynamics of the Romanian equity market are characterised by identifiable regime shifts that can be forecast using flexible nonlinear models. When incorporated into a disciplined allocation framework, these forecasts appear capable of generating economically meaningful improvements in portfolio performance.

References

1. Ang, A. & Bekaert, G. (2002) International asset allocation with regime shifts. *The Review of Financial Studies*, 15(4), pp. 1137–1187. https://doi.org/10.1093/rfs/15.4.1137
2. Ang, A. & Timmermann, A. (2012). Regime changes and financial markets. *Annual Review of Financial Economics*, 4, pp. 313–337. https://doi.org/10.1146/annurev-financial-110311-101808
3. Baum, L.E., Petrie, T., Soules, G. & Weiss, N. (1970). A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *The Annals of Mathematical Statistics*, 41(1), pp. 164–171. https://doi.org/10.1214/aoms/1177697196
4. Bekaert, G. & Harvey, C.R. (1997). Emerging equity market volatility. *Journal of Financial Economics*, 43(1), pp. 29–77. https://doi.org/10.1016/S0304-405X(96)00889-6
5. Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), pp. 307–327. https://doi.org/10.1016/0304-4076(86)90063-1
6. Campbell, J.Y., Lo, A.W. & MacKinlay, A.C. (1997). The Econometrics of Financial Markets. Princeton, NJ: Princeton University Press. https://doi.org/10.1515/9781400830213
7. Chen, T. & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: ACM, pp. 785–794. https://doi.org/10.1145/2939672.2939785
8. Engle, R.F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4), pp. 987–1007. https://doi.org/10.2307/1912773
9. Fama, E.F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2), pp. 383–417. https://doi.org/10.2307/2325486
10. Friedman, J.H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5), pp. 1189–1232. https://doi.org/10.1214/aos/1013203451
11. Gu, S., Kelly, B. & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, 33(5), pp. 2223–2273. https://doi.org/10.1093/rfs/hhaa009
12. Hamilton, J.D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2), pp. 357–384. https://doi.org/10.2307/1912559

13. Hastie, T., Tibshirani, R. & Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd edn. New York: Springer. https://doi.org/10.1007/978-0-387-84858-7
14. J.P. Morgan/Reuters (1996). RiskMetrics—Technical Document. New York: J.P. Morgan.
15. Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1), pp. 77–91. https://doi.org/10.1111/j.1540-6261.1952.tb01525.x
16. Rabiner, L.R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2), pp. 257–286. https://doi.org/10.1109/5.18626