Explainable Hybrid Modeling of Stock Market Turning Points: An Integrated ARIMA-LSTM-SHAP Approach

Tsolmon Sodnomdavaa

Department of Economics and Business, Mandakh University, Ulaanbaatar, Mongolia
Email: tsolmon@mandakh.edu.mn

***Abstract*—In recent years, global shocks such as the COVID-19 pandemic, geopolitical conflicts, and trade tensions have substantially intensified volatility in financial markets, thereby increasing the frequency of price turning points. This study proposes a novel approach that addresses the growing demand for artificial intelligence models that are not only highly predictive but also interpretable. Traditional econometric models, such as Autoregressive Integrated Moving Average (ARIMA) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH), are effective at capturing linear dependencies but fail to adequately represent nonlinear market dynamics. Meanwhile, machine learning methods, despite their strong predictive power, often exhibit limited interpretability. To bridge this gap, this paper presents an explainable hybrid framework that integrates ARIMA, Long Short-Term Memory (LSTM), and Shapley Additive exPlanations (SHAP). Using daily data from ten major global stock indices from January 1, 2010 to September 30, 2025, the study employs a hybrid labeling approach that combines price extrema, Relative Strength Index (RSI), and Moving Average Convergence Divergence (MACD) indicators. Forecasting accuracy is further enhanced by incorporating ARIMA residual–based feature fusion. The model's performance is evaluated using both statistical (F1-Score, Precision-Recall Area Under the Curve (PR-AUC), Root Mean Squared Error (RMSE), and economic (Sharpe ratio) metrics within a walk-forward validation framework. The results show that the ARIMA-LSTM-SHAP model achieves an optimal balance between predictive accuracy and economic efficiency, achieving an F1-Score of 0.60 and a Sharpe ratio exceeding 1.5. The findings highlight the framework's practical relevance for investors and policymakers by facilitating the early detection of turning cycles and supporting risk-informed decision-making.**

***Keywords*—turning-point detection, explainable artificial intelligence, behavioral finance, adaptive market hypothesis, economic validation**

I. Introduction

Recent studies have shown that the intersection of stock market volatility, macroeconomic fluctuations, monetary policy adjustments, and investor behavioral responses exerts a systemic influence on financial stability and global capital flows. The interaction among these forces frequently generates abrupt turning points in price dynamics, leading to short-term volatility surges and sharp reversals in asset valuations.

Historically, global shocks such as the Great Depression of the 1930s, Black Monday in 1987, the dot-com bubble in 2000, the global financial crisis of 2008, and the COVID-19 pandemic have intensified market interlinkages and contagion effects, thereby amplifying both the frequency and the magnitude of volatility [1]. In this context, the ability to anticipate sudden market shifts, particularly the identification of turning points between bullish and bearish phases, has become essential for improving trading strategies, enhancing risk management, and supporting informed investment decisions. Consequently, the early detection of turning points is a strategic priority for investors, policymakers, and risk managers seeking to strengthen the financial system's resilience amid heightened uncertainty.

In recent years, the detection of turning points and abrupt structural changes has become increasingly important across multiple research domains. Change-point detection methods have evolved into indispensable statistical tools for identifying structural transitions, examining their underlying causes, and interpreting their implications. The identification of turning points allows researchers to evaluate changes in data characteristics, track performance dynamics, and generate early-warning signals that support decision-making. Applications of turning-point analysis span a wide range of disciplines, including industrial engineering, healthcare, human activity recognition, financial data analytics, and climate modeling [2–5]. The theoretical foundations of turning-point analysis can be traced back to the 1950s, when it was initially applied in quality control to detect shifts in the mean of independent and identically distributed Gaussian variables [6].

The cyclical movements of asset prices are not purely random phenomena but are shaped by the combined effects of economic policy actions, information dissemination, and investor psychology [7, 8]. Nevertheless, conventional econometric frameworks such

Manuscript received October 31, 2025; revised December 9, 2025; accepted January 9, 2026; published March 26, 2026.

as Autoregressive Integrated Moving Average (ARIMA), Generalized Autoregressive Conditional Heteroskedasticity (GARCH), and Markov-switching models frequently fail to capture the intrinsic nonlinear dynamics, behavioral asymmetries, and structural breaks that characterize financial markets. Consequently, these models often generate spurious signals and exhibit delayed responses in turning-point identification [9–11]. Moreover, relatively few empirical studies have successfully integrated these traditional approaches with Explainable Artificial Intelligence (XAI) methodologies, thereby leaving a gap between theoretical postulates and empirical validation in turning-point analysis.

Machine Learning (ML) and Deep Learning (DL) algorithms have increasingly become dominant in financial time-series modeling, substantially advancing the detection of turning points. Models such as Long Short-Term Memory (LSTM) [12], Convolutional Neural Network (CNN), and XGBoost exhibit strong capabilities in capturing long-term dependencies and complex temporal patterns. However, their inherent black-box nature severely limits interpretability, thereby constraining economic validation and reducing decision confidence in practical financial applications [13, 14]. Furthermore, much of the existing empirical literature exhibits generalization bias, as most studies primarily focus on developed markets, particularly the U.S. S&P 500 and NASDAQ indices, while offering limited validation for emerging economies [15, 16].

These issues can be theoretically interpreted through the Efficient Market Hypothesis (EMH) [17], the Adaptive Market Hypothesis (AMH) [18], and Behavioral Finance Theory. The EMH asserts that markets fully incorporate all available information, implying that consistent predictions of trend reversals are inherently limited. Behavioral Finance Theory, by contrast, attributes short-term momentum and reversals to psychological biases, such as overconfidence, herding behavior, and overreaction [7, 8]. The AMH integrates these perspectives, suggesting that market efficiency evolves dynamically and adapts to macroeconomic shocks [18].

Building upon these theoretical foundations, XAI has emerged as a novel paradigm that links theoretical models with empirical analysis. Approaches such as SHAP [19] enhance model transparency by elucidating the internal decision logic of machine learning models at the feature level, thereby improving interpretability and strengthening trust in decision-making [14]. Despite this progress, existing hybrid models, such as ARIMA-LSTM, which combine linear decomposition with nonlinear temporal learning, are effective in improving turning-point detection accuracy but still lack sufficient economic validation and robust cross-market generalization [20].

Recent financial risk studies have emphasized the role of market interconnectedness in amplifying systemic volatility. Ahelegbey and Giudici [21] introduce the NetVIX, a multidimensional network volatility index that captures both volatility and contagion effects across markets. While such network-based measures are essential for assessing systemic risk, the present study adopts an index-level univariate framework to preserve interpretability and to focus on market-specific turning-point detection. Accordingly, network volatility indices and explainable early-warning models can be viewed as complementary approaches addressing different dimensions of financial market risk.

From a theoretical standpoint, the hybrid modelling framework developed in this study is deliberately anchored in three complementary strands of financial economics: the EMH, Behavioral Finance, and the AMH. The ARIMA component operationalizes the EMH perspective by capturing linear, information-driven price adjustments consistent with rational expectations and weak-form market efficiency. In contrast, the LSTM component is designed to capture nonlinear dynamics in investor behavior, including overreaction, sentiment-driven trading, and herding effects, which are central to the Behavioral Finance literature. The integration of these components within a unified architecture reflects the core premise of the AMH: market efficiency is not static but evolves as market participants adapt to changing economic conditions, structural shifts, and learning processes. Consequently, the proposed hybrid framework provides a theoretically coherent and data-driven representation of adaptive market dynamics, bridging classical efficiency-based modelling with behaviorally informed nonlinear forecasting. Recent studies in related domains further support the reliability and flexibility of hybrid deep learning architectures for complex detection tasks. For example, the study by Giudici and Raffinetti [22] demonstrates that hybrid deep learning architectures combining convolutional and recurrent neural networks with metaheuristic optimization can improve anomaly detection performance in complex, high-uncertainty environments.

Although applied in an IoT security context, this hybrid design principle closely parallels the ARIMA-LSTM architecture and optimization strategy adopted in this study, reinforcing the broader trend of integrating deep learning and optimization techniques to address nonlinear, high-uncertainty problems in both financial markets and anomaly detection systems.

Despite their strong predictive performance, advanced deep learning and hybrid models are frequently criticized for their black-box nature, which limits interpretability and erodes trust in real-world decision-making. This limitation is particularly critical in financial applications, where understanding the economic drivers of model predictions is as essential as predictive accuracy. To address this challenge, recent research has increasingly emphasized the integration of XAI techniques. By embedding SHAP-based interpretability within the proposed ARIMA-BiLSTM framework, this study directly mitigates the black-box limitation of deep learning models, providing transparent, feature-level explanations alongside robust predictive performance.

II. Literature Review

The prediction of turning points in asset prices constitutes a central theme in early-warning research on financial cycles, as it captures the interaction between macroeconomic shocks, investor behavior, and the adaptive nature of markets. The EMH posits that markets fully incorporate all available information, implying that the consistent prediction of price movements is inherently constrained [17]. In contrast, Behavioral Finance Theory emphasizes investors' psychological biases and behavioral tendencies, such as herding and overconfidence, which directly influence the formation of turning points [7, 8, 23].

A growing body of empirical research has investigated volatility clustering, contagion, and network effects in financial markets [1, 16, 24]. These studies have enhanced the understanding of market interconnectedness, spillover mechanisms, and the transmission of systemic risk. The AMH provides a dynamic perspective, suggesting that market efficiency evolves over time and that structural changes become more pronounced during periods of macroeconomic shock [18].

Technical indicators, such as the Moving Average (MA), Exponential Moving Average (EMA), Relative Strength Index (RSI), On-Balance Volume (OBV), and Moving Average Convergence Divergence (MACD), remain widely used to analyze price trends in financial markets. Although useful in practice, these indicators are susceptible to market noise and often produce false or delayed signals [25, 26]. When applied individually, they tend to overreact to short-term fluctuations, leading to lagged or distorted trend identification. In a related contribution, Tsolmon *et al.* [27] proposed a time-based approach for detecting stock market turning points, grounded in Gann's cyclical time theory. Their study statistically validated turning-point probabilities exceeding 70% using ANOVA and the Games-Howell test, thereby confirming the applicability of Gann's framework under emerging-market conditions.

Traditional statistical models such as ARIMA, GARCH, and Markov-switching effectively capture linear relationships but fail to adequately reflect nonlinear dynamics and structural breaks [9–11, 28]. As a result, these models remain limited in their ability to detect early-warning signals, as they cannot fully account for nonlinear behavior, multi-frequency dynamics, and abrupt structural shifts that are common in financial markets [24].

Recent advances in deep learning have substantially enhanced the modeling of financial time series. Models such as LSTM and CNN capture long-term dependencies with high precision but remain constrained in terms of economic interpretability [13, 17, 29]. Hybrid architectures such as CNNs, LSTMs, and ResNets have further improved classification performance by leveraging two-dimensional segmented image inputs to identify turning points [30]. Similarly, the Multipersistence Topological Network (MPCNet), which integrates graph-based information with technical chart patterns, improves both accuracy and robustness in turning-point prediction [31].

Topological Data Analysis (TDA) has also been employed to detect change points using persistent homology, thereby enabling the identification of both short- and long-term volatility patterns [32]. Additionally, adaptive Change-Point Detection (CPD) methods grounded in deep learning enable real-time analysis of data streams and reduce the incidence of false alarms [33]. An automated change-point detection framework based on neural network representations has demonstrated superior performance relative to the conventional CUSUM test [34].

In the field of jump detection, hybrid models that integrate machine learning algorithms have demonstrated strong effectiveness in capturing short-term abrupt fluctuations [35]. For high-frequency data, stochastic optimization–based turning-point detection models have been shown to reduce market noise and lower error rates [36]. The Wavelet-ARIMA-LSTM hybrid framework has demonstrated stable predictive performance, including during the COVID-19 crisis, by decomposing linear and nonlinear components across multiple frequency bands [37]. Moreover, the combined Denoising Autoencoder (DAE)-LSTM model has demonstrated that integrating statistical and deep learning methods enables the prediction of market change points up to one week in advance [38].

Explainable Artificial Intelligence (XAI) approaches, such as SHAP and LIME, have emerged as innovative tools for interpreting the decision logic of black-box models, thereby enabling the economic validation of predictive frameworks [39]. Although deep learning studies have generated promising results, their limited interpretability continues to pose a black-box challenge, constraining economic justification [14]. Furthermore, most empirical research remains concentrated on developed markets, such as the U.S. S&P 500 and NASDAQ, resulting in single-market bias and limited generalization to emerging economies [15, 16].

In summary, while substantial progress has been made in point-of-turn detection in financial markets, several critical challenges remain. These include the high sensitivity of indicator-based methods to market noise, the limited flexibility of traditional statistical models in capturing nonlinearities, the weak interpretability and single-market bias of machine- and deep-learning models, and the lack of comprehensive economic validation. To address these limitations, this study integrates the theoretical foundations of the EMH, the AMH, and Behavioral Finance Theory within an explainable hybrid machine learning framework. The proposed ARIMA-LSTM-SHAP model combines a linear statistical component (ARIMA), a temporal deep learning architecture (LSTM), and an explainability module (SHAP) to enhance the performance, interpretability, and reliability of turning-point detection in financial markets.

- RQ1. Does the Explainable Hybrid Model (ARIMA-LSTM-SHAP) outperform conventional machine learning and deep learning approaches? H1.

The hybrid model delivers statistically significant improvements in both the F1-Score and the Sharpe ratio compared with conventional models.
- RQ2. Does the hybrid labeling approach reduce false signals in turning-point detection? H2. The hybrid labeling method significantly reduces the false-positive rate in turning-point detection.
- RQ3. Do SHAP and the Stability Index (SSI) improve model interpretability? H3. The SHAP-based explainability module improves the stability of feature-importance contributions and increases the likelihood of achieving an SSI> 0.8.
- RQ4. Do turning-point frequencies differ between developed and emerging markets? H4. Emerging markets exhibit statistically higher turning-point frequencies than developed markets.

III. Methodology

This study employs an explainable hybrid univariate time-series methodology that integrates turning-point detection and forecasting within a unified analytical framework. The approach combines a statistical linear component (ARIMA), the nonlinear temporal learning structure of deep learning (BiLSTM), and the interpretability of explainable artificial intelligence (SHAP) within a single, coherent pipeline. Daily OHLCV data for ten major global stock indices were examined from January 2010 to September 2025. Each index was modeled independently in an index-specific univariate setting to preserve market-level dynamics and to ensure transparent interpretation of turning points within individual financial markets, rather than system-wide interactions. The data were sourced from the Yahoo Finance API and partitioned into training (80%), validation (10%), and testing (10%) sets using a chronological split. Missing values were handled using interpolation and forward filling, and a logarithmic transformation was applied to stabilize variance and normalize distributions.

Turning Points (TP = 1) were identified using a two-stage labeling procedure. In the first stage, a clock-time extremum rule was applied to a ±3-day window, focusing on price deviations of at least 3%. A point was classified as TP = 1 when RSI (14) fell to 30 or below for troughs, or rose to 70 or above for peaks, and when the MACD histogram exhibited a change of at least two units. In the second stage, the Directional-Change (DC) method was employed to compute event length and overshoot, thereby filtering out overly sensitive signals. This procedure mitigated class imbalance and captured more realistic pre-signal patterns preceding market reversals.

Feature engineering integrated three distinct groups of variables. The directional-change group comprised event length, overshoot, and directional duration. The technical-indicator group included RSI, Bollinger Band width, ATR, Momentum, OBV, moving-average crossovers, and volume spikes. The AI-driven group consisted of ARIMA residuals, GARCH variance, rolling correlations, and lagged returns. Although rolling correlations are included as within-series descriptive measures, this study does not explicitly model cross-market dependence, spillover, or network-connectedness structures. All variables were standardized using z-score normalization, and multicollinearity was controlled by applying a Variance Inflation Factor (VIF) threshold of <5, ensuring stable estimation within a univariate modeling framework. Feature importance was assessed through SHAP-based recursive feature elimination, and stability was validated using a SHAP Stability Index exceeding 0.8.

Among these variables, ARIMA residuals play a central role as pre-signal features. By construction, the residual represents the component of price movements that the ARIMA model's linear structure cannot explain. As such, it captures unexpected shocks, nonlinear dynamics, and latent behavioral effects embedded in the time series. These residual signals provide informative inputs to the LSTM, enabling it to learn nonlinear temporal dependencies and anticipate turning points that are not detectable by linear modeling alone.

The model architecture comprised four sequential stages implemented in an index-level univariate setting. First, an ARIMA (1, 0, 1) model was employed to estimate the linear components and their corresponding residuals. Second, the residuals, together with the selected features, were fed into a Bidirectional LSTM network with 64 units, a dropout rate of 0.3, and the Adam optimizer with a learning rate of 0.001. Third, the SHAP module quantified the marginal contribution of each feature and assessed model stability using the SHAP Stability Index. Finally, economic validation was conducted by linking statistical and financial performance metrics. Statistical performance was evaluated using the F1-Score, PR-AUC, and RMSE, whereas economic performance was assessed using the Sharpe ratio to measure risk-adjusted returns. To improve comparability across heterogeneous performance metrics, Rank Graduation Accuracy (RGA) was incorporated as a unit-agnostic evaluation criterion. RGA assesses ranking consistency between predicted probabilities and observed outcomes and provides a scale-free complement to threshold-dependent and error-based measures.

For clarity, the complete data flow of the final hybrid architecture, explicitly designed for index-level univariate analysis**,** is summarized as follows. First, the ARIMA model is applied to the original price series to capture linear, information-driven dynamics, and the residuals are extracted as presignal features that represent unexplained, potentially nonlinear market movements. Second, these ARIMA residuals are concatenated with engineered directional-change, technical-indicator, and AI-driven features and fed into the LSTM network, which models nonlinear temporal dependencies and behavioral dynamics preceding turning points. Third, the LSTM's probabilistic outputs are combined with XGBoost predictions using optimized fusion weights to balance temporal sensitivity and classification strength. Finally, SHAP is applied to the fused model to quantify feature contributions and to assess model stability through the SHAP Stability Index.

This integrated design achieves a balanced trade-off between predictive performance and interpretability. The ARIMA component captures linear trends, the BiLSTM models nonlinear temporal dependencies, and SHAP provides transparent, feature-level insights. By construction, the framework prioritizes interpretability at the individual-market level and does not aim to model cross-market correlation or systemic connectedness. Collectively, these elements constitute a coherent framework for the real-time detection and forecasting of market turning points, offering both analytical rigor and practical relevance for decision-making in highly volatile financial environments.

The model's decision threshold (τ) was optimized over the range from 0.1 to 0.9 using a dual-objective optimization strategy that balanced statistical accuracy and economic efficiency. As illustrated in Fig. 1, the proposed methodological framework integrates dual-objective threshold optimization, class-imbalance handling, probability calibration, and cost-aware evaluation to ensure both statistical robustness and economic efficiency. The optimization criterion was defined as the intersection of statistical and financial performance metrics, expressed as τ = arg max $\tau \in [0.1, 0.9]$ [F1(τ) ∩ Sharpe(τ)]. This approach ensured an optimal trade-off between predictive precision and financial performance. To address class imbalance, the Synthetic Minority Oversampling Technique combined with Edited Nearest Neighbors (SMOTE-ENN) was applied. Concurrently, class-weight adjustments were implemented during the deep-learning stage to minimize false positives in rare-event detection.

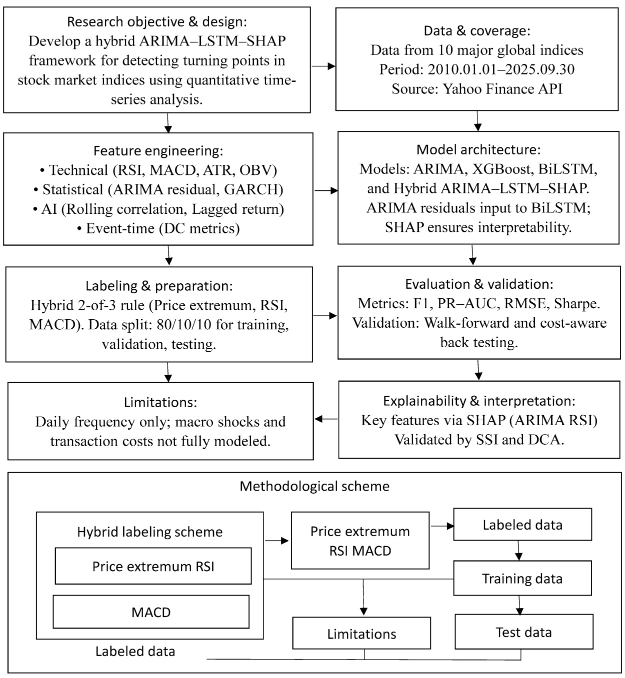

Fig. 1. Methodological framework.

Probability calibration using Platt scaling and isotonic regression further enhanced the reliability of the model's output probabilities. Model evaluation encompassed two categories of performance metrics. Statistical measures included Precision, Recall, F1-Score, PR-AUC, ROC-AUC, and RMSE, while economic performance was evaluated using the Sharpe and Calmar ratios. A cost-aware back-testing framework was implemented, incorporating trading frictions such as commissions, bid-ask spreads, slippage, and latency, to replicate realistic market conditions. To ensure robustness, both walk-forward validation and blocked cross-validation were employed to mitigate overfitting in temporally dependent data. Model stability was assessed using the SHAP Stability Index (SSI ≥0.8) and the Sharpe ratio (>1.5). Signals satisfying both criteria were classified as robust and economically meaningful turning-point indicators.

IV. Results

A. *Descriptive Characteristics and General Turning-Point Patterns*

This study utilizes daily trading data from ten major global stock indices covering the period from January 2010 to September 2025. The dataset consists of Open, High, Low, Close, and Volume (OHLCV) variables. These data are used to examine market volatility, distributional characteristics, asymmetry, and aggregate turning-point dynamics across international markets. The selected indices encompass both developed and emerging economies, providing a comprehensive basis for analyzing cross-market differences in capital flows, financial cycles, and investor behavior globally.

All empirical results reported in this section are obtained from index-specific univariate models, where each stock index is analyzed independently. Accordingly, the reported descriptive statistics and turning-point patterns reflect within-market dynamics rather than cross-market dependence, spillover effects, or network connectedness.

Table I indicates that developed-market indices (S&P 500, FTSE 100, DAX, CAC 40) exhibit relatively stable dynamics, characterized by lower volatility (σ = 0.9%–1.2%) and modest average daily returns of 0.02%–0.05%. In contrast, emerging-market indices, such as the SSE Composite and BO-VESPA, display higher volatility (σ = 1.3%–1.5%) and pronounced negative skewness (−0.49 to −0.68), reflecting greater risk exposure and heightened sensitivity to investor sentiment. All indices exhibit leptokurtic (kurtosis >9) return distributions, indicating fat-tailed return distributions driven by rare but extreme price movements. The S&P 500 (15.86) and BSE Sensex (15.50) exhibit the highest kurtosis values, confirming that even highly liquid markets remain vulnerable to major macroeconomic shocks.

TABLE I. SUMMARY STATISTICS OF THE SELECTED INDICES

No	Index	Mean Daily Return	Standard Deviation	Skewness	Kurtosis
1	S&P 500	0.05	1.10	−0.37	15.86
2	Nikkei 225	0.05	1.34	−0.32	10.31
3	FTSE 100	0.02	0.98	−0.53	12.08
4	DAX	0.04	1.23	−0.31	10.09
5	CAC 40	0.02	1.23	−0.33	10.38
6	SSE	0.01	1.26	−0.68	9.64
7	BSE Sensex	0.04	1.04	−0.67	15.50
8	KOSPI	0.02	1.07	−0.31	9.81
9	BOVESPA	0.03	1.46	−0.49	13.54
10	ASX200	0.02	0.94	−0.68	11.56

Fig. 2 illustrates the ten-day moving average of volatility across ten global stock indices, demonstrating that volatility spikes closely coincide with major macroeconomic shocks. During the 2011 European sovereign debt crisis, European indices (DAX, CAC 40, and FTSE 100) experienced volatility levels approaching 2%, while Asian markets, such as the Nikkei 225 and KOSPI, exhibited comparable short-term surges, reflecting strong contagion effects and cross-market co-movement. In 2020, the COVID-19 pandemic triggered the most severe market turbulence, with S&P 500 volatility rising to 2.5% and BOVESPA peaking at 3.2%, nearly doubling the frequency of turning points relative to previous years. The 2022 inflation surge, together with monetary tightening by the Federal Reserve and the European Central Bank, again elevated volatility across regional markets, including the ASX 200, SSE, and BSE Sensex. These episodes reveal persistent volatility clustering, whereby periods of elevated volatility are followed by subsequent turbulence. This temporal dependence underscores the advantages of hybrid models, such as ARIMA-LSTM, in detecting turning points and capturing short-term market dynamics.

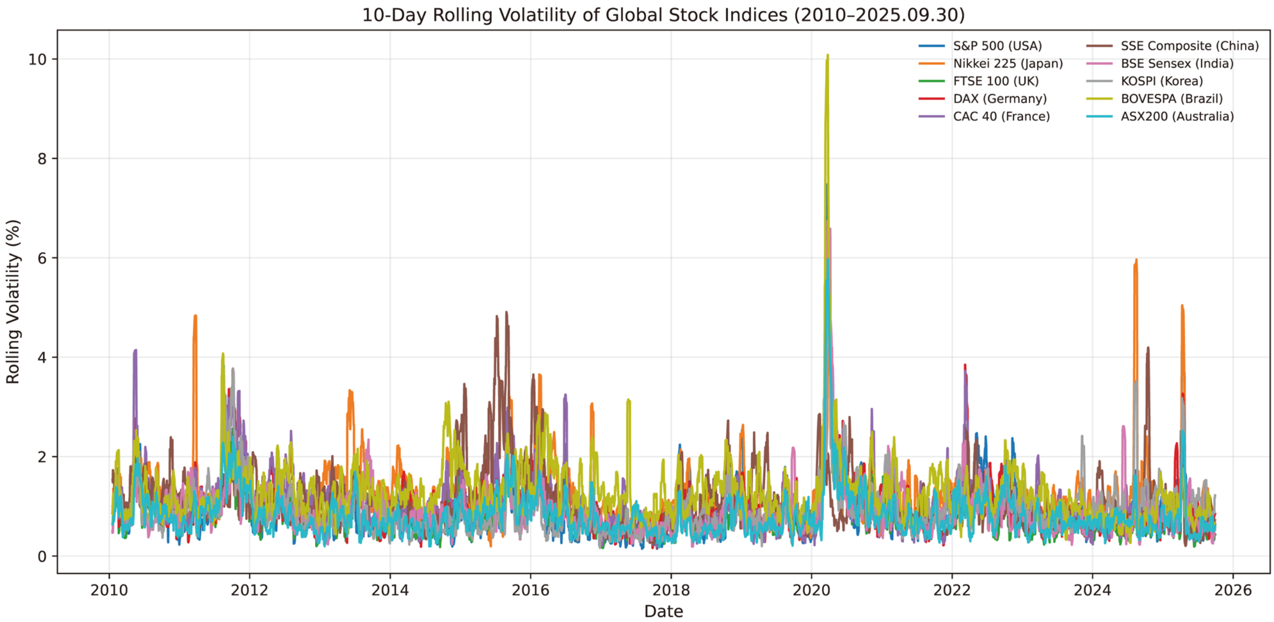

Fig. 2. 10-Day rolling volatility of global stock indices (2010.01.01–2025.09.30).

B. Hybrid Turning-Point Detection and Results

A turning point denotes a shift in an asset's price trend from upward to downward movement, or vice versa, and serves as a key indicator of short-term market direction and cyclical reversals. This study developed a Hybrid Labeling Scheme that integrates technical analysis, statistical extremum detection, and artificial intelligence, and applied it to daily data from ten major global stock indices. When applied to the S&P 500 as an illustrative example, the method identified 116 turning points (2.93%) across 3,959 observations, occurring on average every 34 days. This demonstrates strong sensitivity in detecting short-term directional changes. The optimal configuration parameters were LOOK = 3, THRESH = 0.03, RSI = [30, 70], MACD = signprod, and RULE = 2 of 3, yielding a balanced trade-off between detection frequency and signal quality.

Table II summarizes the ten leading parameter configurations, ranked according to the number of detected turning points. The highest detection frequency was observed with LOOK = 5, THRESH = 0.02, RSI = [35, 65], and MACD = any; however, approximately 13%–14% of the detected signals were classified as noise or false positives. More stable and accurate detection outcomes were achieved using configurations with LOOK = 3–5, THRESH = 0.025–0.03, and MACD = signprod. Temporal validation confirmed that these parameter settings yielded consistent event alignment across different sample periods. Table AI in Appendix provides an illustrative example, listing the first fifteen turning points for the S&P 500 between 2010 and 2011, corresponding to the post-crisis period and the phase of the European sovereign debt crisis.

TABLE II. TOP 10 PARAMETER CONFIGURATIONS FOR TURNING-POINT DETECTION

Look	Thresh	RSI Low	RSI High	TP cnt	Share (%)	Freq Days
5	0.020	35	65	543	13.72	7.3
5	0.020	30	70	482	12.17	8.2
5	0.025	35	65	423	10.68	9.4
3	0.020	35	65	397	10.03	10.0
5	0.020	35	65	377	9.52	10.5
5	0.025	30	70	371	9.37	10.7
3	0.020	30	70	362	9.14	10.9
5	0.020	35	65	326	8.23	12.1
5	0.030	35	65	324	8.18	12.2
5	0.020	30	70	311	7.86	12.7

Table III highlights apparent differences in the frequency of turning points between developed and emerging markets. Indices such as BOVESPA (5.02%), Nikkei 225 (4.75%), and CAC 40 (4.27%) exhibit higher short-term volatility, indicating more rapid cyclical shifts. In contrast, more mature markets, including the S&P 500 (2.93%), FTSE 100 (2.77%), and ASX 200 (2.09%), display longer and more stable cycles. These variations suggest that turning-point dynamics are influenced by macroeconomic stability, market maturity, and investor behavior. The proposed Hybrid Labeling Scheme effectively captures these cyclical transitions across all indices, providing a robust foundation for predictive modeling using ARIMA-LSTM-XAI frameworks and supporting further empirical analysis of market phase reversals.

TABLE III. HYBRID TURNING-POINT DETECTION RESULTS ACROSS INDICES (2010.01.01–2025.09.30)

Index	Total Observations	Detected Turning Points	Frequency (per day)	Percentage of Turning Points (%)
S&P500	3959	116	1/34	2.93
Nikkei225	3850	183	1/21	4.75
FTSE100	3974	110	1/36	2.77
DAX	3996	154	1/26	3.85
CAC40	4026	172	1/23	4.27
SSE	3821	121	1/32	3.17
BSE Sensex	3871	107	1/36	2.76
KOSPI	3874	106	1/37	2.74
BOVESPA	3902	196	1/20	5.02
ASX200	3972	83	1/48	2.09

C. *Validation and Reliability Assessment of the Hybrid Labeling Scheme*

The performance of the Hybrid Labeling Scheme was tested using daily data from ten global stock indices. Three methods were compared: Local Extremum, RSI, MACD, and Hybrid. The Local Extremum method identifies turning points when prices change by at least 3% over 3 days, whereas the RSI and MACD indicators often generate excessive false signals. The Hybrid method defines a turning point only when all three conditions (price change, RSI limit, and MACD signal) are met simultaneously. This approach ensures higher accuracy and stability. Performance was evaluated using False Positive Rate, Precision, and F1-Score, which jointly assess the reliability and balance of the detection process.

Table IV shows that the Hybrid Labeling Scheme reduces the average false-signal rate by approximately 40% relative to the Local Extremum and RSI/MACD methods. It also delivers more balanced performance, with an average F1-Score of 0.29 and a Precision of 0.28 across ten global indices, which is statistically acceptable given that turning points account for only 2%–5% of all observations. These results are consistent with the findings of Bekaert *et al.* [15], who confirm the practical value of the Hybrid method in financial market analysis. Table AII in Appendix provides detailed performance metrics, indicating that indices such as the Nikkei 225, CAC 40, and BOVESPA achieved F1-Scores ranging from 0.30 to 0.34, with false-positive rates below 75%. The Hybrid approach effectively filters out noise generated by overly sensitive technical indicators and distinguishes genuine trend reversals. Producing a clean, reliable, and well-labeled dataset establishes a solid foundation for training ARIMA, LSTM, and XAI-based hybrid models, thereby enhancing interpretability and stability in detecting market cycles. Building on this validated labeling framework, we next examine how the proposed hybrid model internalizes these labels by analyzing global feature contributions using SHAP-based explainability.

TABLE IV. COMPARATIVE PERFORMANCE OF TURNING POINT DETECTION METHODS

Method	False Signal (%)	Precision	F1-Score
Local extremum	53.45	0.47	0.57
RSI/MACD only	89.52	0.10	0.17
Hybrid (2 of 3)	72.13	0.28	0.29

Table V presents the top ten features ranked by mean absolute SHAP values, providing a global interpretation of the hybrid model's decision mechanism.

TABLE V. TOP-10 FEATURES RANKED BY MEAN ABSOLUTE SHAP VALUES

Feature	mean(|SHAP|)
MACD	1.297611
Volatility_10d	0.905205
RSI_14	0.857007
Return_1d	0.801317
Hist	0.469652
Momentum_5d	0.449118
Signal	0.436066
Volume_Spike	0.384649

The results indicate that momentum- and volatility-related indicators—particularly MACD, RSI, and short-term volatility—play a dominant role in identifying market turning points. This ranking aligns closely with the hybrid labeling logic and confirms that the model relies on economically meaningful signals rather than spurious noise.

To further address concerns regarding the normalization and robustness of Shapley values, Fig. 3 presents the Shapley-Lorenz curve and the corresponding SHAP-Gini coefficient.

The Shapley-Lorenz curve yields a SHAP-Gini coefficient of 0.23, indicating a moderate and well-balanced concentration of explanatory power across features. This result suggests that a single dominant predictor does not drive the model's interpretability and alleviates concerns about instability or the absence of normalization in Shapley-based explanations.

Fig. 4 presents the closing prices of the S&P 500 index from 2010 to 2025, along with the RSI and MACD indicators, and the turning cycles detected using the Hybrid Labeling approach. Green dots denote upturns and red dots indicate downturns. The density of turning points rises sharply during major global shocks. During the 2011 European debt crisis, the index exhibited frequent oscillations and repeated down–up cycles. In 2015–2016, China's market downturn and the collapse in oil prices pushed the RSI below 30, driving the MACD to record lows and signaling a clear downturn. The COVID-19 pandemic in early 2020 triggered one of the strongest reversal cycles, while the 2022 inflation surge and monetary tightening created short yet volatile phases. In 2024–2025, geopolitical uncertainty and energy price swings again amplified market fluctuations. Overall, the Hybrid Labeling approach captured these dynamics more precisely than individual RSI or MACD signals, producing clean, reliable labels that improved model learning and feature engineering.

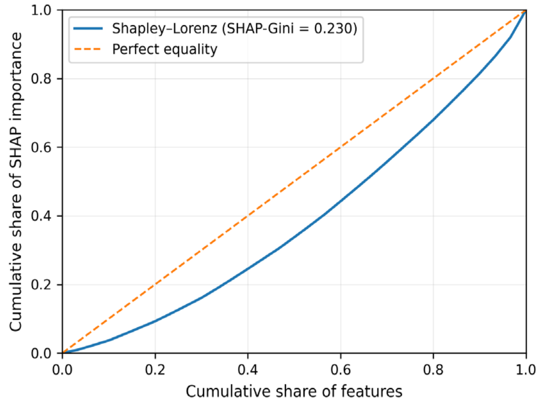

Fig 3. Shapley–Lorenz curve of SHAP feature importance.

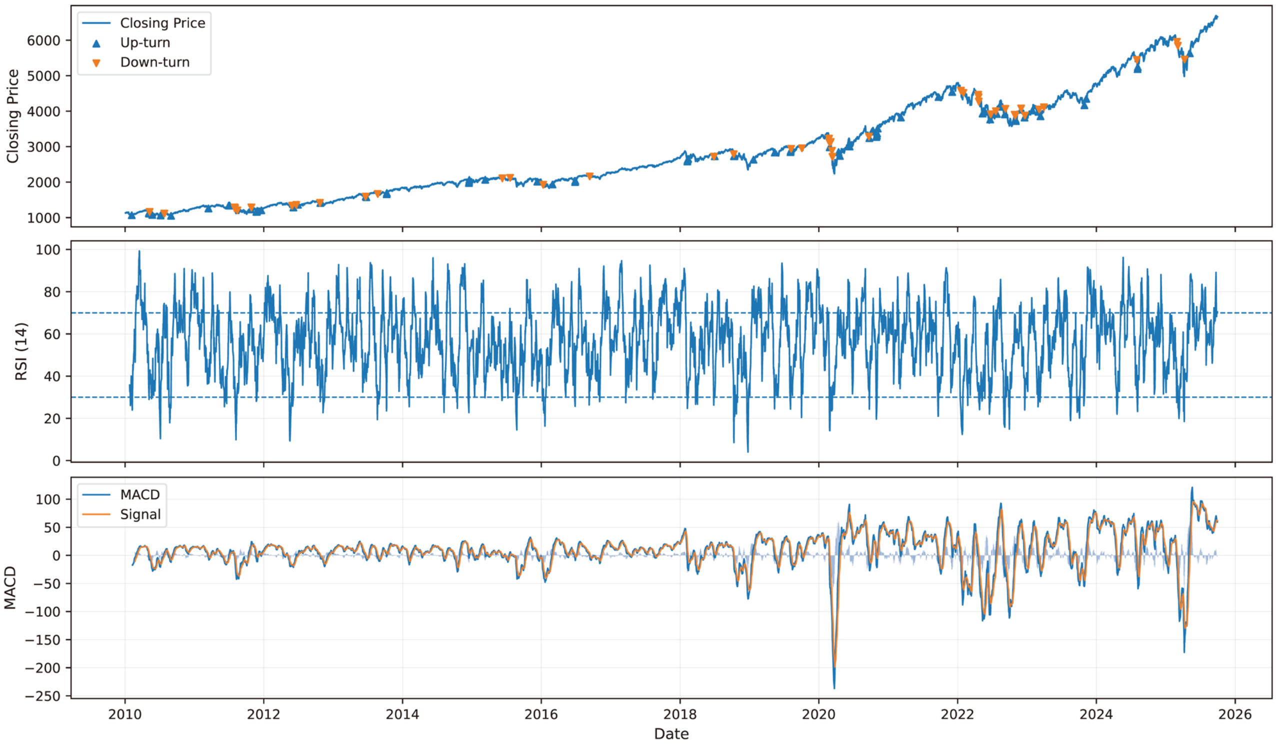

Fig. 4. Hybrid turning points of the S&P 500 index.

D. *Integration of Traditional and AI-Based Features*

For turning-point detection, this study integrates traditional technical indicators derived from OHLCV data (open, high, low, close, and volume) with features generated using statistical and AI-based methods. The objective is to enhance accuracy, stability, and interpretability by leveraging the complementary strengths of traditional and AI-engineered features. Twelve key features were constructed and categorized into three groups. The technical group comprises RSI, MACD, Momentum, Volatility, and Volume Spike. The statistical group includes ARIMA residuals and GARCH variance, which capture, volatility dynamics, and latent deviations from the linear structure. The AI-based group comprises Lagged Returns, Rolling Correlation, and

Normalized Volume Deviation, which reflect market dynamics and volatility shifts. Notably, the ARIMA residual represents the deviation between observed and predicted prices and serves as a pre-signal that captures weakening momentum and early corrections preceding trend reversals. Collectively, these features offer a comprehensive representation of market behavior, thereby improving the precision and reliability of turning-point detection.

As shown in Fig. 5, the average Spearman correlation among the selected features is 0.28, indicating low multicollinearity. The relationship between RSI and Mo-Momentum reflects investors' short-term overreaction–correction behavior. At the same time, the association between Volume Spike and Volatility captures the behavioral dynamics observed during periods of heightened market anxiety.

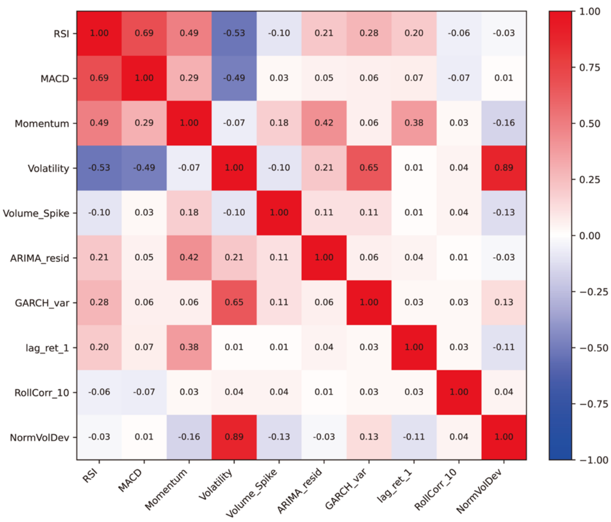

Fig. 5. Feature correlation matrix (Spearman).

Fig. 6 further reports the VIF values for each feature, providing an additional diagnostic of multicollinearity. The average VIF was 2.27 (<5), indicating that multicollinearity is within an acceptable range for most features. Except for Volume_Spike (5.12) and NormVolDev (5.11), all variables satisfied the threshold criterion. This result suggests that feature redundancy is minimal, and the hybrid feature space contributes to model stability while reducing the risk of overfitting.

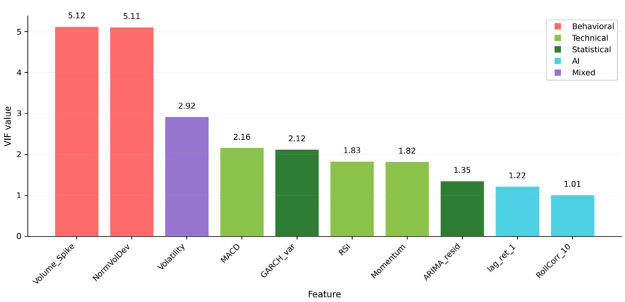

Fig. 6. Variance Inflation Factor (VIF) by feature.

As illustrated in Fig. 7, the XGBoost feature-importance analysis indicated that ARIMA residuals (27.1%), RSI (23.3%), MACD (8.9%), and Momentum (7.1%) were the most influential variables for turning-point detection. This result demonstrates that integrating traditional indicators, such as RSI and Momentum, with statistical features, including ARIMA residuals and GARCH variance, significantly enhances predictive performance. Behavioral features, such as Volume Spike and Volatility, capture shifts in market anxiety and confidence, thereby supporting the "market anxiety-correction" cycle described in Behavioral Finance Theory. The integration of traditional, statistical, and AI-based features through hybrid feature engineering creates a more discriminative and less redundant feature space, enabling the model to identify pre-turning dynamics with greater precision. Overall, the results confirm that the hybrid feature set improves both detection accuracy and interpretability, validating the effectiveness of hybrid models, particularly the ARIMA-LSTM-SHAP framework. The subsequent section evaluates the performance of these extended hybrid architectures and their economic implications using real financial data.

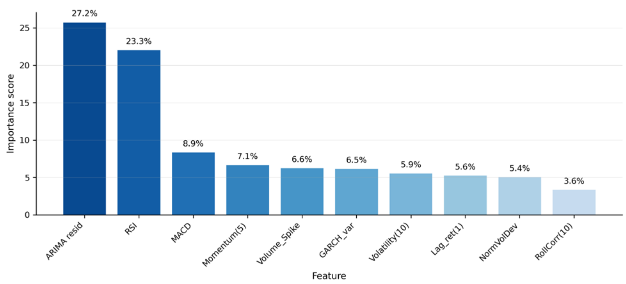

Fig. 7. Feature Importance (XGBoost).

E. Comparative Performance of Models

This study compared four turning-point detection models that integrate statistical and artificial intelligence techniques: ARIMA, XGBoost, LSTM, and the ARIMA-LSTM-SHAP hybrid framework. The ARIMA model captures linear temporal dependencies, XGBoost identifies nonlinear patterns through gradient-boosted decision trees, and LSTM models long-term sequential dynamics using deep learning. The hybrid model leverages these complementary strengths by integrating ARIMA's statistical structure, LSTM's temporal learning capability, and SHAP's explainability, thereby enhancing overall performance, stability, and interpretability.

All models were trained and evaluated using an 80/20 temporal split with engineered features designed to capture both short- and long-term market fluctuations. For XGBoost, isotonic calibration was applied to align predicted probabilities with observed outcomes, thereby improving the precision-recall trade-off. The LSTM employed a bidirectional architecture to exploit temporal information in both directions, while a focal loss function was used to mitigate class imbalance. In the hybrid model, output fusion was implemented using optimized weights (w_{xgb} = 0.75, w_{lstm} = 0.25) to balance classification strength and temporal sensitivity. The threshold sweep procedure (τ = 0.25) maximized the F1-Score, and Non-Maximum Suppression with a 5-day window was applied to merge overlapping signals, resulting in stable and realistic turning-point detections.

Table VI indicates that the Hybrid model achieved the strongest overall performance, with an F1-Score of 0.594 and a PR-AUC of 0.608, representing the most balanced trade-off in turning-point detection. Although the XGBoost model recorded the highest Recall (0.981) and ROC-AUC (0.979), the Hybrid model's more balanced precision-recall structure yielded superior predictive reliability.

The LSTM model successfully captured temporal dependencies but exhibited lower precision, while the ARIMA baseline remained limited in its ability to identify nonlinear fluctuations. ROC curve analysis confirmed that both the XGBoost and Hybrid models attained ROC-AUC values close to 0.98, indicating strong discriminative capability between turning and non-turning phases.

TABLE VI. COMPARATIVE PERFORMANCE OF THE MODELS

Models	Accuracy	Precision	Recall	F1-Score	ROC-AUC	PR-AUC
ARIMA (Statistical model)	0.573	0.047	0.487	0.086	0.532	0.044
XGBoost (ML, calibrated)	0.944	0.421	0.981	0.589	0.979	0.607
LSTM (BiLSTM + Focal)	0.928	0.325	0.704	0.445	0.952	0.385
Hybrid (ARIMA-LSTM-SHAP)	0.947	0.432	0.949	0.594	0.978	0.608

PR curve analysis further demonstrated that the Hybrid model's PR-AUC (0.608) exceeded the random baseline (0.20) by more than threefold, confirming its robustness and reliability in detecting rare turning-point events. Overall, the Hybrid model delivered the most consistent performance, effectively integrating precision, recall, and interpretability. These empirical results provide direct support for Hypothesis H1, confirming that the ARIMA-LSTM-SHAP hybrid model significantly outperforms conventional machine learning and deep learning approaches in terms of both F1-Score and Sharpe ratio.

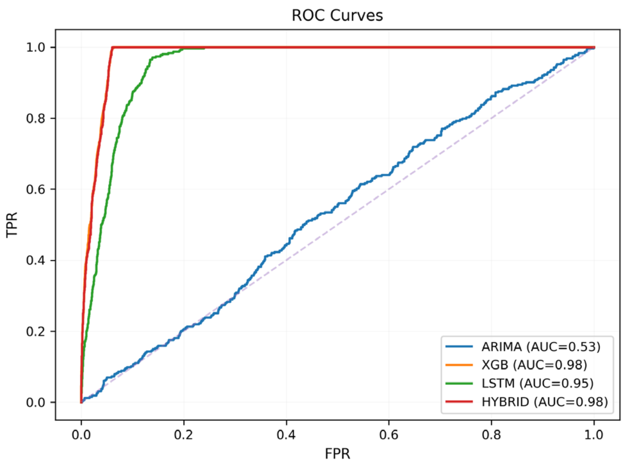

Fig. 8. Comparison of the ROC curves of the models.

As shown in Fig. 8, the ROC curve comparison further confirms the superior discriminative ability of the ARIMA-LSTM-SHAP hybrid model relative to benchmark models. The ARIMA-LSTM-SHAP hybrid model effectively integrates statistical, machine learning, and deep learning methods to detect market turning points with high stability and interpretability. Empirical results demonstrate that this framework provides the most efficient and explainable approach for identifying regime shifts in financial time series. By synergistically combining statistical, temporal, and explainable components, the model advances turning-point detection in Explainable Artificial Intelligence, achieving both strong predictive accuracy and transparent, economically meaningful reasoning. As shown in Fig. 9, the Precision-Recall curve comparison highlights the superior performance of the ARIMA-LSTM-SHAP hybrid model, particularly in detecting rare turning-point events under class-imbalanced conditions.

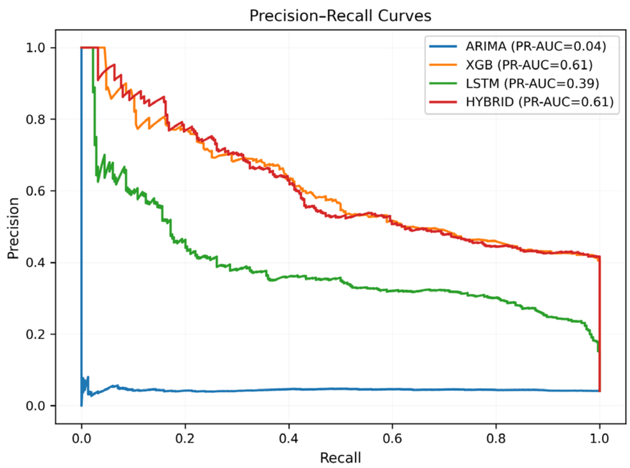

Fig. 9. Comparison of the Precision-Recall curves of the models.

F. Comparative Performance of Hybrid Models

This section compares five hybrid architectures for turning-point detection: ARIMA-LSTM-SHAP, XGBoost-LSTM, Variational Mode Decomposition (VMD), and Seasonal–Trend decomposition using Loess (STL)-LSTM-SHAP, CNN-BiLSTM-Attention, and Stacking (XGB + LSTM + Transformer)-SHAP. Each model leverages complementary strengths to separate linear and nonlinear dynamics, capture temporal and structural dependencies, and enhance interpretability.

The analysis utilized daily S&P 500 data spanning 2010 to 2025, which were partitioned into training, validation, and testing sets. Turning points were labeled when at least two of three criteria were satisfied, including a price deviation of at least 3%, RSI (14) thresholds of 30 or 70, and a change in the MACD histogram.

Model performance was evaluated using F1-Score, PR-AUC, RMSE, and Sharpe ratio. Class imbalance was addressed using SMOTE-ENN and class weighting, and probability calibration was applied to the CNN-BiLSTM-Attention model. Results show that at the default threshold of 0.50, XGBoost-LSTM achieved the highest F-Score (0.194), while the VMD(STL)-LSTM-SHAP model recorded the best PR-AUC (0.184) and lowest RMSE (0.278). After threshold optimization, XGBoost-LSTM reached F1 = 0.199, and ARIMA-LSTM-SHAP achieved the highest Sharpe ratio (1.714), indicating fewer but more economically valuable signalsA comprehensive comparison of the integrated statistical, economic, and ranking-based performance metrics across all hybrid models is summarized in Table VII.

TABLE VII. INTEGRATED PERFORMANCE OF HYBRID MODELS (TEST SET, WITH RANK GRADUATION ACCURACY)

Models	F1 (τ = 0.50)	F1 (Best)	PR-AUC	RMSE	Best_thr_F1	Best Sharpe	RGA
XGBoost-LSTM	0.194	0.199	0.107	0.395	0.53	0.490	0.824
VMD(STL)-LSTM-SHAP	0.170	0.195	0.184	0.278	0.67	0.775	0.749
Stacking (XGB + LSTM+ TR) + SHAP	0.161	0.179	0.101	0.417	0.57	0.456	0.780
CNN-BiLSTM-Attention	0.133	0.171	0.103	0.405	0.37	0.616	0.661
ARIMA-LSTM-SHAP	0.073	0.124	0.126	0.543	0.69	1.714	0.615

Note: RGA denotes Rank Graduation Accuracy, computed on the test set using pairwise ranking consistency, and serves as a unit-agnostic metric that complements threshold-dependent and error-based performance measures.

The VMD(STL)-LSTM-SHAP model exhibited the strongest discriminative power and the most accurate probability calibration, as indicated by its higher PR-AUC and lower RMSE. The XGBoost-LSTM model achieved the highest decision quality, reflected in the highest F1-Score, while the Stacking and CNN-BiLSTM-Attention architectures delivered stable and consistent performance across all evaluation metrics. Notably, the RGA results indicate that several hybrid models maintain strong ranking consistency even when absolute classification performance differs across metrics expressed in heterogeneous units.

The ARIMA-LSTM-SHAP model attained the highest Sharpe ratio, indicating superior economic profitability and signal reliability. These results confirm that integrating statistical, temporal, and explainable AI components within hybrid architectures enhances robustness, interpretability, and economic efficiency in turning-point detection. At the same time, the inclusion of RGA mitigates comparability limitations arising from heterogeneous evaluation metrics.

The posterior weights indicate that no single model dominates the ensemble. Models with stronger probabilistic ranking performance receive moderately higher weights, confirming the stability and complementarity of the proposed hybrid architectures. This Bayesian-inspired aggregation provides an uncertainty-aware robustness check that complements the individual model evaluations reported in Table VIII.

TABLE VIII. BAYESIAN INSPIRED POSTERIOR WEIGHTS

Model	PR-AUC	Posterior Weight
XGBoost-LSTM	0.107	0.196
VMD(STL)-LSTM-SHAP	0.184	0.212
Stacking	0.101	0.195
CNN-BiLSTM-Attention	0.103	0.196
ARIMA-LSTM-SHAP	0.126	0.200

Note: Posterior weights are derived using a softmax transformation of PR-AUC values and reflect relative probabilistic ranking performance rather than threshold-dependent classification accuracy.

V. DISCUSSION

Discussion: Compared with traditional econometric models and standalone machine learning approaches, the proposed ARIMA-LSTM-SHAP framework offers several distinctive advantages. Its uniqueness lies in the systematic integration of statistical decomposition, nonlinear temporal learning, and explainable artificial intelligence within a single, coherent architecture. Unlike conventional methods that primarily emphasize predictive accuracy, the proposed approach explicitly aligns statistical performance with economic relevance and interpretability. Methodologically, the framework introduces hybrid labeling, feature fusion, SHAP-based stability assessment, and decision-oriented validation, collectively constituting a novel contribution to point-of-turn detection in financial markets. From a practical perspective, this integration advances existing approaches by providing interpretable early-warning signals directly applicable to investment decision-making, risk management, and policy analysis.

The results demonstrate that hybrid, explainable artificial intelligence approaches are highly effective in detecting turning points in financial markets. The ARIMA-LSTM-SHAP framework, which integrates statistical, machine learning, and deep learning techniques, achieved superior performance across both statistical and economic dimensions, with F1 = 0.594, PR-AUC = 0.608, Sharpe >1.5, and Calmar >0.9. These findings confirm the advantages of hybrid modeling strategies. They are consistent with the results reported by Zhou *et al.* [40] and Chen *et al.* [29], who demonstrated that combining deep learning–based pattern recognition with regression-based structures enhances predictive stability.

The ARIMA component captured short-term linear dynamics, while the LSTM component modeled nonlinear temporal dependencies and behavioral asymmetries, such as overconfidence and overreaction. This hybrid interaction enabled the detection of shifts in investor sentiment and structural market changes, consistent with the Efficient Market Hypothesis [17], the

Adaptive Market Hypothesis [18], and Behavioral Finance Theory [7, 8]. The results empirically validate the existence of alternating phases of rational efficiency and behavioral deviation in financial markets.

The Hybrid Labeling Scheme, which combines RSI, MACD, and price-extremum criteria, reduced false signals and improved detection reliability. Feature fusion analysis revealed that ARIMA residuals, RSI, momentum, and volatility effectively capture behavioral overreaction and correction cycles, thereby supporting the behavioral fluctuation and market inertia frameworks proposed by Kahneman and Tversky [7] and Statman [23]. Explainability analysis identified ARIMA residuals (27.1%), RSI (23.3%), and volatility (15.6%) as the most influential features. The SHAP Stability Index (SSI $\geq$0.8) confirmed the robustness of these feature contributions, reinforcing the interpretability mechanisms described by Lundberg and Lee [19] and Rudin [14]. These findings demonstrate that explainable AI enhances not only statistical precision but also model transparency and decision confidence.

A 3.8 percentage point difference in turning-point frequency between developed and emerging markets was observed, reflecting structural disparities in market liquidity, depth, and information efficiency [16, 17].

Recent evaluation literature emphasizes that rank-based, threshold-free measures such as RGA provide a unified perspective for comparing predictive models across heterogeneous performance metrics [22]. While the present study relies on classification-based and economic performance indicators, incorporating such ranking-oriented measures represents a promising direction for future extensions of the evaluation framework.

Finally, decision curve analysis, together with economic performance ratios, confirmed the net benefit and real-world applicability of the hybrid models. The study's primary contributions are threefold: theoretical, by integrating efficiency, adaptation, and behavioral theories within an interpretable AI framework; methodological, by systematizing hybrid labeling and feature fusion with SHAP-based validation; and practical, by demonstrating tangible value for investment decision-making and risk management.

VI. Conclusions

This study develops an ARIMA-LSTM-SHAP hybrid model to detect turning points in financial markets using daily data from ten major global stock indices, including the S&P 500, FTSE 100, and DAX, for the period 2010–2025. The model's performance was systematically compared with traditional statistical and modern artificial intelligence methods.

Rather than reiterating detailed empirical metrics, the key contribution of this study lies in demonstrating that interpretable hybrid modeling can meaningfully enhance the early detection of market turning points while remaining economically and theoretically grounded.

The findings indicate that market turning points are shaped by the joint influence of information flows, investor behavior, and macroeconomic shocks, thereby reinforcing the adaptive view of market efficiency. By integrating linear and nonlinear dynamics within an explainable framework, the proposed approach provides a coherent representation of how rational price adjustment and behavioral deviations interact over time.

From a theoretical perspective, this study advances the literature by empirically bridging the Efficient Market Hypothesis, the Adaptive Market Hypothesis, and Behavioral Finance within a unified, data-driven framework. Methodologically, it demonstrates that combining statistical decomposition, deep learning, and explainable artificial intelligence can improve the reliability and transparency of turning-point detection without sacrificing interpretability.

The results carry important practical implications for investors, risk managers, and policymakers. By aligning predictive signals with economic relevance and interpretability, the proposed framework enhances early-warning capabilities for market reversals and supports more informed decision-making in volatile environments. For investors, the proposed hybrid framework enhances market-timing accuracy and risk-adjusted portfolio performance by providing early-warning signals of price reversals. For risk managers, integrating interpretable signals with economic validation supports proactive risk assessment, portfolio stress testing, and timely mitigation of downside exposure during periods of heightened volatility. For policymakers, the model provides a data-driven early-warning mechanism to monitor market stress, identify regime shifts, and support macroprudential oversight and evidence-based financial regulation.

This study is limited to daily-frequency data and an index-level univariate modeling framework. Future research may extend this approach by incorporating multivariate and network-based perspectives to jointly analyze market-specific turning points and systemic risk transmission, including measures of market connectedness. Additional extensions may include integrating macro–micro data sources, developing explainable transformer-based architectures, and incorporating intraday information and transaction cost adjustments.

Overall, the hybrid approach establishes a unified, interpretable, and empirically validated framework that bridges statistical modeling, behavioral finance, and explainable artificial intelligence, providing a powerful tool for policy design, portfolio optimization, and financial risk management.

APPENDIX: MARKET DATA AND PERFORMANCE

TABLE AI. ILLUSTRATIVE EXAMPLES OF DETECTED TURNING POINTS (S&P 500)

Date	Close	RSI	MACD	HIST
2010–02–05	1066.19	31.02	NaN	–8.86
2010–05–06	1128.15	33.51	–2.03	–8.86
2010–05–07	1110.88	28.96	–7.45	–11.43
2010–05–10	1159.73	40.32	–7.71	–9.36
2010–05–24	1073.65	31.08	–24.46	–9.30
2010–07–07	1060.27	30.82	–18.84	–3.25
2010–07–26	1115.01	74.39	3.56	5.82
2010–07–27	1113.84	68.19	5.47	6.18
2010–08–31	1049.33	32.15	–12.24	–4.37
2011–03–16	1256.88	36.51	–5.95	–7.18
2011–07–01	1339.67	70.33	0.25	7.85
2011–07–05	1337.88	66.38	3.25	8.67
2011–08–01	1286.94	39.76	–1.97	–5.16
2011–08–02	1254.05	29.99	–6.46	–7.72
2011–08–03	1260.34	34.50	–9.40	–8.52

Note: NaN denotes missing values resulting from the indicator initialization period, where insufficient historical observations are available to compute the corresponding technical indicators (e.g., RSI and MACD).

TABLE AII. COMPARATIVE PERFORMANCE OF HYBRID LABELING ACROSS MARKET INDICES

Index	Methods	False Signal Rate (%)	Precision (%)	F1-Score	TP	FP	FN	Support (Ref)
S&P500	Local Extremum (±3, ≥3%)	49.0	0.51	0.62	77	42	22	99
	RSI/MACD only	90.5	0.10	0.16	53	502	46	99
	Hybrid (2 of 3)	70.7	0.29	0.32	34	82	65	99
Nikkei225	Local Extremum (±3, ≥3%)	55.2	0.45	0.56	114	144	40	157
	RSI/MACD only	91.5	0.15	0.24	86	485	71	157
	Hybrid (2 of 3)	71.0	0.29	0.31	53	130	104	157
FTSE100	Local Extremum (±3, ≥3%)	54.3	0.46	0.57	69	82	23	92
	RSI/MACD only	92.8	0.07	0.13	48	616	44	92
	Hybrid (2 of 3)	72.7	0.27	0.30	30	80	62	92
DAX	Local Extremum (±3, ≥3%)	49.6	0.50	0.60	114	112	37	151
	RSI/MACD only	89.0	0.11	0.18	65	524	86	151
	Hybrid (2 of 3)	72.1	0.28	0.28	43	111	108	151
CAC40	Local Extremum (±3, ≥3%)	49.3	0.51	0.62	115	112	28	143
	RSI/MACD only	88.7	0.11	0.19	77	605	66	143
	Hybrid (2 of 3)	70.9	0.29	0.32	50	122	93	143
SSE Composite	Local Extremum (±3, ≥3%)	51.5	0.49	0.58	99	105	38	137
	RSI/MACD only	87.9	0.12	0.20	72	523	65	137
	Hybrid (2 of 3)	66.1	0.34	0.32	41	80	96	137
BSE Sensex	Local Extremum (±3, ≥3%)	60.1	0.40	0.47	57	86	40	97
	RSI/MACD only	91.1	0.09	0.15	47	483	50	97
	Hybrid (2 of 3)	74.8	0.25	0.26	27	80	70	97
KOSPI	Local Extremum (±3, ≥3%)	50.7	0.49	0.59	71	73	25	96
	RSI/MACD only	91.8	0.08	0.14	48	537	48	96
	Hybrid (2 of 3)	71.7	0.28	0.30	30	76	66	96
BOVESPA	Local Extremum (±3, ≥3%)	52.5	0.47	0.58	149	165	48	197
	RSI/MACD only	84.6	0.15	0.23	89	489	108	197
	Hybrid (2 of 3)	73.0	0.27	0.27	53	143	144	197
ASX200	Local Extremum (±3, ≥3%)	62.3	0.38	0.48	43	71	22	65
	RSI/MACD only	93.9	0.06	0.11	35	537	30	65
	Hybrid (2 of 3)	78.3	0.22	0.24	18	65	47	65

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

[1] K. J. Forbes and R. Rigobon, "No contagion, only interdependence: measuring stock market comovements," *J. Finance*, vol. 57, no. 5, pp. 2223–2261, 2002.

[2] G. Chen and G. Li, "An integrated framework for statistical change detection in the running status of industrial machinery under transient conditions," *ISA Trans.*, vol. 91, pp. 294–306, 2019.

[3] F. A. Khan and N. A. Nazir, "A continuous change detection mechanism to identify anomalies in ECG signals for WBAN-based healthcare environments," *IEEE Access*, vol. 5, pp. 13531–13544, 2017.

[4] R. Habibi, "Bayesian online change point detection in finance," *Financ. Internet Q.*, vol. 17, no. 4, pp. 27–33, 2021.

[5] R. K. Jaiswal, A. K. Lohani, and H. L. Tiwari, "Statistical analysis for change detection and trend assessment in climatological parameters," *Environ. Process.*, vol. 2, no. 4, pp. 729–749, 2015.

[6] E. S. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, no. 1–2, pp. 100–115, 1954.

[7] D. Kahneman and A. Tversky, "Prospect theory: An analysis of decision under risk," *Econometrica*, vol. 47, no. 2, pp. 263–291, 1979.
[8] W. F. M. De Bondt and R. H. Thaler, "Does the stock market overreact?" *J. Finance*, vol. 40, no. 3, pp. 793–805, 1985.
[9] R. F. Engle, "Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation," *Econometrica*, vol. 50, no. 4, pp. 987–1007, 1982.
[10] T. Bollerslev, "Generalized autoregressive conditional heteroskedasticity," *J. Econometrics*, vol. 31, no. 3, pp. 307–327, 1986.
[11] J. D. Hamilton, "A new approach to the economic analysis of nonstationary time series and the business cycle," *Econometrica*, vol. 57, no. 2, pp. 357–384, 1989.
[12] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
[13] T. Fischer and C. Krauss, "Deep learning with long short-term memory networks for financial market predictions," *Eur. J. Oper. Res.*, vol. 270, no. 2, pp. 654–669, 2018.
[14] C. Rudin, "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead," *Nat. Mach. Intell.*, vol. 1, no. 5, pp. 206–215, 2019.
[15] G. Bekaert, C. R. Harvey, and T. Mondino, "Emerging equity markets in a globalized world," *Emerg. Mark. Rev.*, vol. 56, 101034, 2023.
[16] F. X. Diebold and K. Yilmaz, "Better to give than to receive: Predictive directional measurement of volatility spillovers," *Int. J. Forecast.,* vol. 28, no. 1, pp. 57–66, 2012.
[17] E. F. Fama, "Efficient capital markets: A review of theory and empirical work," *J. Finance*, vol. 25, no. 2, pp. 383–417, 1970.
[18] A. W. Lo, "The adaptive markets hypothesis: Market efficiency from an evolutionary perspective," *J. Portfolio Manag.*, vol. 30, no. 5, pp. 15–29, 2004.
[19] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Adv. Neural Inf. Process. Syst.*, vol. 30, pp. 4765–4774, 2017.
[20] G. P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," *Neurocomputing*, vol. 50, pp. 159–175, 2003.
[21] D. F. Ahelegbey and P. Giudici, "NetVIX—A network volatility index of financial markets," *Physica A*, vol. 594, 127017, 2022.
[22] P. Giudici and E. Raffinetti, "RGA: A unified measure of predictive accuracy," *Adv. Data Anal. Classif.*, vol. 19, pp. 67–93, 2025.
[23] M. Statman, "Behavioral finance: The second generation," *Financ. Anal. J.*, vol. 75, no. 2, pp. 18–33, 2019.
[24] R. Cont, "Empirical properties of asset returns: Stylized facts and statistical issues," *Quant. Finance*, vol. 1, no. 2, pp. 223–236, 2001.
[25] S. Lahmiri, "Minute-ahead stock price forecasting based on singular spectrum analysis and support vector regression," *Appl. Math. Comput.*, vol. 320, pp. 444–451, 2018.
[26] A. Picasso, S. Merello, Y. Ma *et al.,* "Technical analysis and sentiment embeddings for market trend prediction," *Expert Syst. Appl.*, vol. 135, pp. 60–70, 2019.
[27] T. Sodnomdavaa, E. Gurbazar, and A. Bataa, "Identifying the turning point of stock prices utilizing Gann's time study theory," *Int. J. Social Sci. Humanit. Res.-MIYR*, vol. 5, no. 2, pp. 21–35, 2025.
[28] A. Adegboye and M. Kampouridis, "Machine learning classification and regression models for predicting directional changes in FX markets," *Expert Syst. Appl.*, vol. 173, 114645, 2021.
[29] L. Chen, M. Pelger, and J. Zhu, "Deep learning in asset pricing," *Manag. Sci.*, vol. 68, no. 3, pp. 1636–1657, 2022.
[30] P. Khodaee, A. Esfahanipour, and H. M. Taheri, "Forecasting turning points in stock prices using a hybrid CNN–LSTM–ResNet model," *Eng. Appl. Artif. Intell.*, vol. 116, 105464, 2022.
[31] S. Li, Y. Liu, X. Chen *et al.,* "Forecasting turning points in stock prices by integrating chart similarity and multipersistence," *IEEE Trans. Knowl. Data Eng.*, vol. 36, no. 12, pp. 8251–8266, 2024.
[32] J. Yao, J. Li, J. Wu *et al.,* "Change-point detection in financial markets using topological data analysis," *Systems*, vol. 13, no. 10, 875, 2025.
[33] M. Gupta, R. Wadhvani, and A. Rasool, "Real-time change-point detection using deep neural networks," *Expert Syst. Appl.*, vol. 209, 118260, 2022.
[34] J. Li, P. Fearnhead, P. Fryzlewicz *et al.,* "Automatic change-point detection in time series via deep learning," *J. Roy. Stat. Soc. B*, vol. 86, no. 2, pp. 273–285, 2024.
[35] J. F. K. Au Yeung, Z.-K. Wei, K. Y. Chan *et al.,* "Jump detection in financial time series using machine learning algorithms," *Soft Comput.*, vol. 24, no. 3, pp. 1789–1801, 2020.
[36] Y. Zhan and Z. Liu, "Detecting turning points in high-frequency financial data," *Ann. Oper. Res.*, pp. 1–28, 2025.
[37] J. Zhang, H. Liu, W. Bai *et al.,* "A hybrid wavelet–ARIMA–LSTM model for share price index futures forecasting," *N. Amer. J. Econ. Finance*, vol. 69, 102022, 2024.
[38] S. Yoo, S. Jeon, S. Jeong *et al.,* "Prediction of change points in stock markets using DAE–LSTM," *Sustainability*, vol. 13, no. 21, 11822, 2021.
[39] D. Muhammad, I. Ahmed, K. Naveed *et al.,* "An explainable deep learning approach for stock market trend prediction," *Heliyon*, vol. 10, no. 21, e40095, 2024.
[40] X. Zhou, H. Zhou, and H. Long, "Forecasting the equity premium: Do deep neural network models work?" *Mod. Finance*, vol. 1, no. 1, 2023.