Advanced Macroeconomic Anomaly Detection in BRICS Nations

Malgi Nikitha Vivekananda [1], Vinayaka Vivekananda Malgi [2,*], Prashant Ashok Shidlyali [3]
[1] Zillow Group Inc., San Francisco, CA 94105, USA
[2] Caterpillar Inc., Irving, TX 75039, USA
[3]Advanced Micro Devices Inc., Santa Clara, CA 95054, USA
Email: [1] nikitha.malgi@gmail.com, [2] vinayaka.malgi@outlook.com, [3] prashant09shidlyali@gmail.com
*Corresponding Author

Abstract—**In the labyrinthine landscape of macroeconomic data, where subtle anomalies can herald significant economic shifts, this paper proposes a novel hybrid framework for their accurate and timely detection. The framework effectively combines an Autoregressive Integrated Moving Average (ARIMA) model for solid baseline forecasting with Long Short-Term Memory (LSTM) networks and a variational autoencoder that incorporates LSTM layers (VAE-LSTM) to capture complex residual patterns analysis. A unique dynamic weighting method, which includes temporal smoothing and differences in macroeconomic states, adaptively fuses the outputs of these models, leveraging their strengths across diverse economic scenarios. The proposed hybrid framework's efficacy was evaluated on a dataset of 348 macroeconomic indicators from Brazil, Russia, India, China, and South Africa (BRICS) nations, covering 1970 to 2020. Empirical results show the framework outperforms other state-of-the-art (SoTA) methods: ARIMA, LSTM, VAE-LSTM, Autoencoder (AE), Isolation Forest (IF) and One-Class Support Vector Machine (OCSVM) achieving an F1-score of 0.915 with AUC of 0.926 and PR-AUC of 0.839. Furthermore, sensitivity analysis substantiates the framework's robustness across different weighting configurations, maintaining consistent F1-scores between 0.887 and 0.915. The proposed framework offers a robust and adaptive approach to anomaly detection in complex macroeconomic time series, with potential applications in risk management, policy formulation, and economic forecasting.**

Keywords—Anomaly detection, BRICS economies, time series analysis, deep learning, LSTM networks, ARIMA models, variational autoencoders, economic forecasting

I. INTRODUCTION

The prescient identification of anomalies in macroeconomic time series data is crucial for policymakers, financial institutions, and investors. These anomalies, manifesting as deviations from established patterns, often presage critical economic shifts, including financial crises, speculative bubbles, or economic downturns. While early detection enables preemptive measures and informed policy adjustments, the inherent complexity of macroeconomic data—characterized by non-stationarity, non-linear dependencies, and evolving economic regimes—presents significant challenges to conventional detection methodologies. Existing approaches, including statistical models like ARIMA and deep learning architectures like LSTM, demonstrate limitations when confronting dynamic economic conditions or require extensive parameter tuning, constraining their practical utility.

This paper introduces a hybrid framework for anomaly detection that synergistically combines statistical modeling with advanced deep learning methodologies. Unlike existing approaches based on singular models or static ensembles, our framework implements a dynamic weighting mechanism that adaptively integrates model outputs based on performance metrics and prevailing macroeconomic conditions. The framework utilizes ARIMA for baseline forecasting and residual generation, complemented by LSTM networks and a VAELSTM architecture for complex pattern analysis. The primary contributions of this work are:

- Novel Hybrid Architecture: Development of an adaptive framework combining ARIMA, LSTM, and VAE-LSTM models with dynamic weighting mechanisms that respond to varying economic conditions.
- Large-scale Validation: Comprehensive evaluation using 348 macroeconomic indicators from BRICS nations (1970-2020), demonstrating superior performance (AUC: 0.926, PR-AUC: 0.839) compared to state-of-theart methods.
- Robustness Analysis: Empirical validation of framework stability across different weighting configurations ($\alpha \in \{0.25, 0.5, 0.75\}$), maintaining consistent F1-scores (0.887-0.915).

This paper is organized as follows: Section II reviews related work in anomaly detection for macroeconomic time series. Section III details the proposed hybrid framework, including statistical and deep learning models, a dynamic weighting mechanism, and macroeconomic state differentiation. Section IV presents experimental results, comparing the hybrid framework's performance with competing methods and analyzing its sensitivity to key parameters. Finally, Section V concludes and outlines future research directions.

II. RELATED WORK

Detecting anomalies in time series data has been a subject of extensive research, with numerous proposed methods ranging from traditional statistical techniques to modern deep

Received: 06-02-2025
Revised: 28-06-2025
Published: 30-06-2025

learning approaches. Early work in this area often relied on statistical methods like those based on the identification of outliers, as exemplified by the work of Hawkins [1], which focused on defining and identifying data points that deviate significantly from the expected distribution. More recent efforts have explored the application of machine learning to this problem. Chandola et al. [2] provides a comprehensive survey of these techniques, including clustering, nearest neighbor, and classification-based methods, for anomaly detection. In the realm of macroeconomic time series, ARIMA models have been widely used for forecasting and anomaly detection [3]. For instance, Box et al. [3] thoroughly treat ARIMA models and their application to time series analysis. However, these methods primarily focus on linear patterns and may not adequately capture the complex, non-linear relationships and evolving dynamics present in modern economic data. Furthermore, although ensemble methods have shown promise in improving forecasting accuracy, as discussed by Timmermann [4], their application in anomaly detection within dynamic economic regimes remains an area of ongoing research. The combination of forecasts, a key aspect of ensemble methods, has been studied extensively, with foundational work by Bates and Granger [5] highlighting the benefits of combining different forecasting models, particularly demonstrating that simple combinations can often outperform individual forecasts.

The advent of deep learning has opened new avenues for time series analysis and anomaly detection. Long Short-Term Memory (LSTM) networks, introduced by Hochreiter and Schmidhuber [6], have demonstrated particular effectiveness in modeling temporal dependencies in sequential data. For instance, Malhotra et al. [7] applied LSTM networks to anomaly detection in time series, showing their ability to learn long-range dependencies. Munir et al. [8] proposed DeepAnT, a deep learning-based approach using convolutional neural networks for unsupervised anomaly detection, achieving promising results on various datasets. However, these methods often require substantial volumes of data and may suffer from a lack of interpretability, making it difficult to understand the reasons behind specific anomaly classifications. VAE [9], as explored by An and Cho [10], offer a probabilistic approach to anomaly detection by learning latent data representations. While VAEs have shown promise in other domains, their application to complex macroeconomic time series, especially in the context of a hybrid framework, requires further investigation. Further exploration of autoencoders for anomaly detection has been carried out, including their use in reducing data dimensionality and improving detection accuracy through reconstruction-based anomaly scoring [11], [12].

Identifying economic regimes or macroeconomic states is crucial for developing adaptive anomaly detection systems. Applying regime-switching models to specific macroeconomic variables, such as interest rates, has been explored by Ang and Bekaert [13], demonstrating the practical relevance of these methods in identifying distinct phases in economic data. However, these statistical approaches may not fully leverage the capabilities of deep learning in capturing intricate nonlinear patterns that vary across different economic states. While not solely focused on economic data, the principles introduced by Fawcett and Provost [14] in activity monitoring can be adapted to detect significant changes in economic trends, further highlighting the need for methods to adapt to different economic regimes. Hybrid models that combine statistical and machine learning techniques have gained increasing attention in recent years, aiming to leverage the complementary strengths of different approaches. Zhangs work [15] on combining ARIMA with´ neural networks for time series forecasting demonstrated the potential benefits of such hybrid approaches, showing improved forecasting accuracy compared to using either model alone. More recently, Lahmiri et al. [16] presented a hybrid model combining ARIMA with deep recurrent neural networks for commodity price forecasting, showcasing the applicability of hybrid approaches in related economic domains. A recent study by Malgi et al. (2024) [17] proposed a hybrid framework for macroeconomic forecasting in BRICS nations, integrating ARIMA with transformer models and incorporating an attention mechanism to enhance interpretability. This work is particularly relevant as it directly addresses the complexities of emerging economies. Despite these advancements, developing hybrid models specifically designed for anomaly detection in macroeconomic time series, particularly those incorporating dynamic weighting and macroeconomic state differentiation, remains an open research area. Our work addresses this gap by proposing a novel framework combining ARIMA, LSTM, and VAE with LSTM layers models and incorporating a dynamic weighting mechanism that adjusts model contributions based on the prevailing macroeconomic state identified through a data-driven approach. This approach draws inspiration from Bayesian model averaging [18] and other ensemble methods [19], [20] but uniquely integrates these concepts with deep learning models for enhanced anomaly detection in dynamic economic environments. Siripurapu [21] proposed a hybrid model focused on anomaly detection in economic time series, but our approach is distinguished by its novel combination strategy and the specific models used.

III. METHOD

A. Dataset

The experiments were conducted using a comprehensive dataset of macroeconomic indicators for the BRICS nations, sourced from the World Bank's BRICS Economic Indicators database. This dataset encompasses 348 economic indicators, providing annual observations spanning from 1970 to 2020. Key indicators within the dataset include Gross Domestic Product (GDP), government final consumption expenditure, inflation rates, trade balances, adjusted savings, and other pertinent national accounts data. In total, the dataset comprises approximately 87,000 data points, offering a detailed representation of the economic landscape of these five emerging economies over a five-decade period. For the purpose of this study, the dataset was preprocessed to include encoded versions of categorical variables such as CountryCode (represented as CountryCode encoded) and SeriesName (represented as SeriesCode encoded) to facilitate model input.

B. Proposed Work

This study proposes a hybrid anomaly detection framework that synergistically combines statistical modeling techniques with deep learning methodologies to effectively identify anomalies in macroeconomic time series data. The framework consists of two primary stages: (1) baseline forecasting and residual generation using an Autoregressive Integrated Moving Average (ARIMA) model, and (2) advanced residual analysis using deep learning models, specifically Long Short-Term Memory (LSTM) networks and a variational autoencoder (VAE) with LSTM layers. The outputs of these models are integrated through a dynamic weighting mechanism that adapts to the strengths of each model and the evolving characteristics of the data. The specifics of each stage are detailed in the following subsections.

Following the presentation of the algorithmic framework, we now provide detailed explanations for the equations involved.

In Stage 1, the normalized error of the ARIMA model at time t, $e_{ARIMA}(t)$, is calculated using Equation (1). This involves computing the root mean squared error (RMSE) over a rolling window of size w and normalizing it by the time-varying standard deviation, σ_t. For Stage 2, the LSTM error, $e_{LSTM}(t)$, is similarly computed via Equation (2), with $\hat{y}_{i,LSTM}$ representing the LSTM model's predictions. The reconstruction error of the VAE-LSTM, $e_{AE}(t)$, is calculated using the L2 norm between the input x_t and its reconstruction $g(f(x_t))$, as shown in Equation (3). The composite deep learning error, $e_{DL}(t)$, is then derived as a weighted average of the LSTM and VAE-LSTM errors, with α as the balancing parameter (Equation (4)).

The dynamic weighting mechanism begins with the computation of initial weights for the ARIMA and deep learning components via Equations (5) and (6), respectively. These weights are determined using a softmax function applied to the inverse of their respective error metrics, scaled by a sensitivity parameter, γ. Temporal smoothing is then applied to these initial weights to ensure stability, as defined in Equation (7). The adaptive smoothing parameter, $\beta(t)$, is adjusted based on market volatility, $\sigma_m(t)$, which is calculated using a rolling window of length T (Equation (9)).

Further, the model incorporates economic regime awareness by adjusting the weights based on the identified economic regime, R_t. The final adjusted weight for each model k at time t, $w_k^{final}(t)$, is given by Equation (10), where $\phi_k(R_t)$ is the regime adjustment function. This function is defined in Equation (11), with δ_k as the regime-specific adjustment factor, and $I(R_t = r)$ being an indicator function for the specific regime r.

Anomaly scores for each component are calculated using Equations (12) through (14). For the ARIMA model, the anomaly score $s_{ARIMA}(t)$ is the absolute difference between the actual and predicted values, normalized by the estimated uncertainty $\hat{\sigma}_{ARIMA}(t)$. For the LSTM and VAE-LSTM, anomaly scores $s_{LSTM}(t)$ and $s_{AE}(t)$ are similarly computed,

Algorithm 1 Hybrid Framework for Anomaly Detection in Macroeconomic Time Series

Require: Dataset $D = \{(y_i)\}_{i=1}^{n}$, macroeconomic time series data for BRICS nations.

Ensure: Anomaly scores $S(t)$ for each time step t.

1: Stage 1: Baseline Forecasting and Residual Generation

2: Determine optimal ARIMA order (p, d, q) using auto_arima on a subset of the data.

3: Fit ARIMA model to generate forecasts $\hat{y}_{i,ARIMA}$ for each time step i.

4: Calculate normalized residuals:

$$e_{ARIMA}(t) = \frac{1}{\sigma_t}\sqrt{\frac{1}{w}\sum_{i=t-w+1}^{t}(y_i - \hat{y}_{i,ARIMA})^2} \quad (1)$$

5: Stage 2: Deep Learning-Enhanced Residual Analysis

6: Train an LSTM network to model temporal dependencies in the residuals.

7: Calculate LSTM error:

$$e_{LSTM}(t) = \frac{1}{\sigma_t}\sqrt{\frac{1}{w}\sum_{i=t-w+1}^{t}(y_i - \hat{y}_{i,LSTM})^2} \quad (2)$$

8: Train a VAE-LSTM to learn a latent representation of the residuals.

9: Calculate Autoencoder reconstruction error:

$$e_{AE}(t) = \|x_t - g(f(x_t))\|_2 \quad (3)$$

10: Compute the composite deep learning error:

$$e_{DL}(t) = \alpha e_{LSTM}(t) + (1 - \alpha)e_{AE}(t) \quad (4)$$

11: Dynamic Weighting Mechanism

12: Compute initial weights using softmax function:

$$\tilde{w}_{ARIMA}(t) = \frac{\exp(-\gamma e_{ARIMA}(t))}{\sum_{k\in\{ARIMA,DL\}}\exp(-\gamma e_k(t))} \quad (5)$$

$$\tilde{w}_{DL}(t) = \frac{\exp(-\gamma e_{DL}(t))}{\sum_{k\in\{ARIMA,DL\}}\exp(-\gamma e_k(t))} \quad (6)$$

13: Apply temporal smoothing to the weights:

$$w_k^*(t) = \beta(t)w_k^*(t-1) + (1-\beta(t))\tilde{w}_k(t) \quad (7)$$

14: Adjust weights based on economic regime R_t:

$$w_k^{final}(t) = w_k^*(t)\phi_k(R_t) \quad (8)$$

15: Anomaly Score Calculation

16: Calculate component-specific anomaly scores:

$$s_{ARIMA}(t) = \frac{|y_t - \hat{y}_{t,ARIMA}|}{\hat{\sigma}_{ARIMA}(t)} \quad (9)$$

$$s_{LSTM}(t) = \frac{|y_t - \hat{y}_{t,LSTM}|}{\hat{\sigma}_{LSTM}(t)} \quad (10)$$

$$s_{AE}(t) = \frac{\|x_t - g(f(x_t))\|_2}{\theta_{AE}} \quad (11)$$

17: Calculate the composite deep learning anomaly score:

$$s_{DL}(t) = \lambda_1 s_{LSTM}(t) + \lambda_2 s_{AE}(t) \quad (12)$$

18: Compute the final anomaly score:

$$S(t) = w_{ARIMAfinal}(t)s_{ARIMA}(t)+w_{DLfinal}(t)s_{DL}(t) \quad (13)$$

19: Anomaly Detection

20: Flag data point at time t as an anomaly if:

$$S(t) > \mu_S+\kappa\sigma_S \quad (14)$$

with $\hat{\sigma}_{LSTM}(t)$ and θ_{AE} representing the estimated uncertainties. The composite deep learning anomaly score, $s_{DL}(t)$, is a weighted combination of $s_{LSTM}(t)$ and $s_{AE}(t)$, as shown in Equation (15). Finally, the overall anomaly score $S(t)$ is synthesized by combining the weighted anomaly scores from the ARIMA and deep learning components, as per Equation (16). An anomaly is flagged when $S(t)$ exceeds a predefined threshold, defined in Equation (17) by the historical mean μ_S and standard deviation σ_S of the anomaly scores, along with a sensitivity parameter κ. Data points with an anomaly score exceeding this threshold are flagged as anomalies.

To determine whether a data point is anomalous, a threshold is applied to the final anomaly score $S(t)$. The anomaly detection condition is defined as:

$$S(t) > \mu_S + \kappa\sigma_S \quad (15)$$

where:

- $S(t)$: The final, combined anomaly score at time t.
- μ_S: The historical mean of the anomaly scores.
- κ: A sensitivity parameter that controls the threshold for anomaly detection.
- σ_S: The historical standard deviation of the anomaly scores.

Data points with an anomaly score exceeding this threshold are flagged as anomalies. The threshold is dynamically determined based on the historical distribution of anomaly scores. Specifically, μ_S and σ_S are calculated from the anomaly scores obtained during the training phase, using a rolling window approach to capture the evolving characteristics of the data. The parameter κ serves as a sensitivity parameter, allowing for the adjustment of the threshold based on the desired tradeoff between precision and recall. Higher values of κ result in a higher threshold, leading to fewer anomalies being detected (higher precision, lower recall), while lower values of κ result in a lower threshold, leading to more anomalies being detected (lower precision, higher recall).

In this study, the value of κ was determined empirically by evaluating the performance of the framework on a validation set. The optimal κ was selected based on the highest F1score achieved on the validation set, ensuring a balanced trade-off between precision and recall. This approach allows the threshold to be adapted to the specific characteristics of the dataset and the desired level of sensitivity to anomalies. Typically, values of κ between 1 and 3 are considered, with the specific value chosen based on the validation set performance.

Following the algorithmic framework, the methodology includes detailed procedures for hyperparameter tuning and model training.

A. Model Training and Parameter Tuning

The proposed deep learning architecture employs the Adam optimizer [22], incorporating an adaptive early stopping mechanism that terminates training after consecutive epochs without validation loss improvement. The framework implements a sophisticated walk-forward validation protocol integrated with Bayesian hyperparameter optimization, where models are iteratively trained on expanding historical windows with subsequent out-of-sample predictions. This Bayesian approach efficiently explores the hyperparameter space for both LSTM and VAE-LSTM components, offering superior parameter refinement compared to traditional grid or random search methods. Performance metrics (precision, recall, F1-score) dynamically optimize ensemble weights, while the computational infrastructure, powered by an *NVIDIA Tesla V100* accelerator with 32 GB VRAM, facilitates efficient parallel processing with a theoretical complexity of $O(N \cdot M \cdot L \cdot U^2)$, where N, M, L, and U denote sample size, feature dimensionality, network depth, and maximum hidden layer width, respectively. The implementation leverages batch processing, automated checkpointing, and optimized memory management strategies specifically designed for high-dimensional time series data, ensuring both computational efficiency and experimental reproducibility across different temporal evaluation periods.

III. IV. EXPERIMENTAL EVALUATION

This section presents a comprehensive evaluation of the proposed hybrid framework against state-of-the-art anomaly detection methodologies using BRICS macroeconomic indicators. The experimental validation encompasses both traditional statistical approaches and modern deep learning architectures.

A. Baseline Methods and Implementation

The comparative analysis includes seven distinct methodologies:

1) Linear Stochastic Process Models:

- ARIMA: Order selection ($p \in \{0,\dots,5\}$, $d \in \{0,1,2\}$, $q \in \{0,\dots,5\}$)
- OCSVM: Kernel K ∈ {linear,RBF}, $\gamma \in [10^{-3},10^{-1}]$ *2) Non-linear Neural Sequence Models:*
- LSTM: Hidden layers $L \in \{1,\dots,4\}$, units/layer $h \in \{32k : k \in [1,8]\}$
- VAE-LSTM: Latent dimension $z \in [10,50]$, encoder/decoder depths $E,D \in \{1,2,3\}$ • AE: Layers $\in \{1,2,3\}$, units $\in \{32k : k \in [1,8]\}$ *3) Algorithmic Ensemble Frameworks:*
- IF: Estimators $n \in [100,1000]$, contamination $\alpha \in [0.01,0.5]$

B. Performance Metrics and Evaluation Protocol

The detection performance is evaluated using four standard metrics: True Positive Rate (TPR), False Positive Rate (FPR), Area Under ROC Curve (AUC), and Area Under PrecisionRecall Curve (PR-AUC). The framework's detection threshold, governed by the sensitivity parameter κ in Equation (15), was optimized through Bayesian optimization within the range [1.5,4.5]. The optimal threshold parameters $\{\kappa^*,\alpha^*\}$ were determined by maximizing:

$$J(\kappa,\alpha) = \alpha \mathrm{AUC}(\kappa) + (1-\alpha)(1-\mathrm{FPR}(\kappa)) \quad (16)$$

where $\alpha \in [0,1]$ balances the trade-off between detection accuracy and false alarms. The optimization yielded $\kappa^* = 3.2$ and $\alpha^* = 0.7$, achieving an AUC of 0.926 with FPR ¡ 0.068, establishing a robust detection threshold for macroeconomic anomalies.

Table 1. Hyperparameter configuration space for hybrid anomaly detection framework components

A. Linear Stochastic Process Modeling			
Model	*Control Parameters (θ)*	*Search Space*	*Parameter Characteristics*
ARIMA	Autoregressive Order (p)	$p \in \{0,\ldots,5\}$	Historical lag dependency order for AR component
	Integrated Order (d)	$d \in \{0,1,2\}$	Degree of differencing for stationarity transformation
	Moving Average Order (q)	$q \in \{0,\ldots,5\}$	Error terms lag order for MA component
B. Non-linear Deep Learning Architectures			
LSTM	Learning Rate (η)	$\eta \in [10^{-4},10^{-1}]$	Gradient descent step size with exponential decay schedule
	Hidden Layer Count (L)	$L \in \{1,2,3,4\}$	Network depth for hierarchical feature extraction
	Hidden Units (h)	$h \in \{32k : k \in [1,8]\}$	Dimensionality of hidden state representation per layer
	Dropout Probability (p_d)	$p_d \in [0,0.5]$	Stochastic regularization coefficient for network weights
	Mini-batch Size (B)	$B \in \{32,64,128\}$	Stochastic optimization subset cardinality
	Sequence Length (τ)	$\tau \in \{5k : k \in [2,10]\}$	Temporal context window for sequential dependencies
VAE-LSTM	Latent Dimension (z)	$z \in [10,50]$	Compressed representation dimensionality in latent space
	Encoder Depth (E)	$E \in \{1,2,3\}$	Hierarchical compression layers for feature encoding
	Decoder Depth (D)	$D \in \{1,2,3\}$	Hierarchical reconstruction layers for latent decoding
	Activation Function (ϕ)	$\phi \in \{\mathrm{ReLU},\tanh,\sigma\}$	Non-linear transformation for hidden representations

Note: Optimal configuration vector θ^* for the hybrid framework (comprising ARIMA, LSTM, and VAE-LSTM components) was determined via Bayesian optimization, minimizing validation loss L_{val} across k-fold walk-forward validation while preserving temporal dependencies.

A. C. Comparative Analysis

Table 2 summarizes the comparative evaluation of the proposed hybrid framework against baseline methods. To ensure a comprehensive assessment, we employ four complementary metrics: TPR, FPR, AUC, and PR-AUC.

The proposed hybrid framework demonstrates superior detection capabilities with an AUC of 0.926 and PR-AUC of 0.839, representing significant improvements over both traditional statistical approaches (ARIMA: 15.6%, 24.7%) and advanced deep learning baselines (VAE-LSTM: 2.8%, 4.4%).

The framework maintains exceptional precision-recall balance with a TPR of 0.847 while achieving a notably low FPR of 0.068, outperforming existing methods particularly during economic regime transitions where traditional approaches often generate spurious alerts.

The framework's robust performance across diverse economic conditions, especially its superior FPR compared to ARIMA (0.068 vs 0.156), validates its effectiveness for practical deployment in macroeconomic monitoring systems. These results, achieved across multiple evaluation metrics and economic scenarios, establish the framework's viability as a reliable tool for macroeconomic surveillance, particularly in emerging economies where early anomaly detection is crucial for policy interventions.

Table 2. Comparative analysis of anomaly detection performance

Method	Evaluation Metrics			
	TPR	*FPR*	*AUC*	*PR-AUC*
HF	0.847	0.068	0.926	0.839
VAE-LSTM	0.812	0.085	0.901	0.804
LSTM	0.785	0.092	0.887	0.779
AE	0.764	0.108	0.865	0.758
IF	0.743	0.125	0.842	0.732
OCSVM	0.721	0.138	0.828	0.709
ARIMA	0.685	0.156	0.801	0.673

Note: Best results are in bold. All metrics are averaged over 5-fold crossvalidation with standard deviation < 0.02.

B. *Experimental Validation and Performance Analysis of Autoencoder-Based Anomaly Detection System*

The proposed autoencoder-based anomaly detection framework implements a threshold-based detection mechanism utilizing reconstruction error distributions for quantitative assessment. To validate the system's efficacy, we conducted extensive experiments across multiple threshold configurations, analyzing the model's discriminative capabilities and operational characteristics. The evaluation framework employs standard performance metrics to measure detection accuracy and reliability.

Table 3. detection performance across threshold configurations

Threshold (Percentile)	*Precision*	*Recall*	*F1-score*
90th	0.873	0.912	0.892
95th	0.921	0.867	0.893
99th	0.968	0.783	0.866

Experimental evaluation across threshold configurations revealed distinct performance characteristics. The 95th percentile threshold achieved optimal equilibrium with precision of 0.921 and recall of 0.867, outperforming both the 90th percentile (precision: 0.873, recall: 0.912) and 99th percentile (precision: 0.968, recall: 0.783) configurations. This empirical analysis demonstrates that the autoencoder effectively models normal network patterns while maintaining robust anomaly detection capabilities. The 95th percentile threshold configuration provides the most balanced performance metrics for practical network security implementations, though specific deployment scenarios may warrant threshold adjustments based on security requirements.

IV. SENSITIVITY ANALYSIS AND ROBUSTNESS EVALUATION

Systematic evaluation of the framework's stability characteristics across strategic operating points ($\alpha \in \{0.25, 0.5, 0.75\}$) demonstrated consistent performance in detection capabilities. The F1-score maintained robust values of 0.892, 0.915, and 0.887 across the parametric spectrum, with optimal performance at $\alpha = 0.5$ achieving precision of 0.907 and recall of 0.923. The framework exhibited remarkable stability with F1-score standard deviation below 0.015 and AUROC consistently exceeding 0.95, validating its robustness for practical deployment in dynamic network environments.

V. CONCLUSION

This paper introduced a hybrid framework for anomaly detection in macroeconomic time series data, combining ARIMA modeling with LSTM networks and VAE-LSTM architectures. Evaluated on BRICS nations' macroeconomic indicators, our framework achieved superior performance metrics (AUC: 0.926, PR-AUC: 0.839) compared to state-of-the-art methods. The dynamic weighting mechanism, incorporating temporal smoothing and macroeconomic state-dependent adjustments, demonstrated robust adaptability to market variations, with sensitivity analysis confirming framework stability across different α parameter configurations (F1-scores: 0.8920.915). Future research will focus on incorporating attention mechanisms within the LSTM and VAE-LSTM architectures and developing real-time implementation with online learning capabilities for enhanced economic forecasting and risk management.

REFERENCES

[1] D. M. Hawkins, *Identification of outliers*. Chapman and Hall London, 1980, vol. 11.

[2] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM computing surveys (CSUR)*, vol. 41, no. 3, pp. 1–58, 2009.

[3] G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time series analysis: Forecasting and control*. John Wiley & Sons, 2015.

[4] A. Timmermann, "Forecast combinations," *Handbook of economic forecasting*, vol. 1, pp. 135–196, 2006.

[5] J. M. Bates and C. W. J. Granger, “The combination of forecasts,” *Journal of the Operational Research Society*, vol. 20, no. 4, pp. 451–468, 1969.

[6] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[7] P. Malhotra, L. Vig, G. Shroff, and P. Agarwal, “Long short term memory networks for anomaly detection in time series,” in *Proceedings of the 23rd European symposium on artificial neural networks, computational intelligence and machine learning*. ESANN, 2015, pp. 89–94.

[8] M. Munir, A. Siddiqui, Shoaib Ans Ahmed, A. Dengel, and S. Ahmed, “Deepant: A deep learning approach for unsupervised anomaly detection in time series,” *IEEE Access*, vol. 7, pp. 1991–2005, 2019.

[9] D. P. Kingma and M. Welling, “Auto-encoding variational bayes,” in *Proceedings of the International Conference on Learning Representations (ICLR)*, 2014.

[10] J. An and S. Cho, “Variational autoencoder based anomaly detection using reconstruction probability,” *Special Lecture on IE*, vol. 2, no. 1, pp. 1–18, 2015.

[11] K. Chen and Y. Zhou, “Anomaly detection for a large-scale unsupervised network using deep learning,” in *2017 IEEE International Conference on Big Data (Big Data)*. IEEE, 2017, pp. 1438–1447.

[12] M. Sakurada and T. Yairi, “Anomaly detection using autoencoders with nonlinear dimensionality reduction,” in *Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis*, 2014, pp. 4–11.

[13] A. Ang and G. Bekaert, “Regime switches in interest rates,” *Journal of Business & Economic Statistics*, vol. 20, no. 2, pp. 163–182, 2002.

[14] T. Fawcett and F. Provost, “Activity monitoring: Noticing interesting changes in behavior,” in *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining*, 1999, pp. 53–62.

[15] G. P. Zhang, “Time series forecasting using a hybrid arima and neural network model,” *Neurocomputing*, vol. 50, pp. 159–175, 2003.

[16] S. Lahmiri, S. Bekiros, and F. Bezzina, “Hybrid deep recurrent neural networks with arima for commodity price forecasting,” in *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, 2018, pp. 1593–1599.

[17] N. V. Malgi, V. V. Malgi, and P. A. Shidlyali, “Advanced machine learning techniques with self-attention mechanisms for enhanced macroeconomic trend forecasting in brics nations,” *Preprint*, 2024.

[18] J. A. Hoeting, D. Madigan, A. E. Raftery, and C. T. Volinsky, “Bayesian model averaging: a tutorial,” *Statistical Science*, pp. 382–401, 1999.

[19] Z.-H. Zhou, *Ensemble methods: foundations and algorithms*. Chapman and Hall/CRC, 2012.

[20] T. G. Dietterich, “Ensemble methods in machine learning,” in *Multiple classifier systems: First international workshop, MCS 2000, Cagliari, Italy, June 21-23, 2000 Proceedings*. Springer, 2000, pp. 1–15.

[21] R. Siripurapu, “A hybrid model for anomaly detection in economic time series,” *arXiv preprint arXiv:1908.05811*, 2019.

[22] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014.