Advanced International Journal for Research (AIJFR)
E-ISSN: 3048-7641 ● Website: www.aijfr.com ● Email: editor@aijfr.com

Explaining Trade Volatility: A Hybrid Machine Learning and SHAP-Based Interpretability Analysis of India's Foreign Trade (1990–2025)

Balasubramanian S[1], Natarajan M[2]

[1,2] Department of Computer and Information Science, Faculty of Science, Annamalai University, Annamalai Nagar, Chidambaram, Tamil Nadu, India.

DOI: https://doi.org/10.63363/aijfr.2026.v07i06.8731

Abstract

Accurate trade forecasting and structural volatility attribution are vital for macroeconomic stability, tariff policy formulation, and central bank foreign reserve management. However, empirical trade modeling often struggles with trade-offs between linear interpretability and complex predictive capacity. This paper presents an empirical framework evaluating machine learning models alongside Explainable Artificial Intelligence (xAI) techniques to model monthly aggregate exports in India from April 1990 to March 2025 (420 observations). We construct a feature space comprising autoregressive lags, moving averages, rolling volatility metrics, trade balance ratios, and exogenous crisis indicators. Benchmark evaluations across Ridge Regression, Random Forest, Extra Trees, LightGBM, and XGBoost demonstrate that Ridge Regression achieves superior generalization on out-of-sample test data (2020–2025), yielding a Root Mean Squared Error (RMSE) of $3.6592 billion, a Mean Absolute Error (MAE) of $2.6889 billion, a Mean Absolute Percentage Error (MAPE) of 10.05%, and an R^2 score of 0.6278. Non-linear ensemble methods suffer from trend extrapolation bias due to non-stationary structural shifts across the 35-year horizon. Model attribution using Shapley Additive exPlanations (SHAP) reveals that medium-to-long-term trend indicators (Exports_MA_12, Imports_MA_12) and short-term momentum (Exports_Lag_2, Exports_Lag_6) dominate global predictions, while short-term rolling volatility indicators exert localized non-linear pressure during post-shock trade recovery phases.

Keywords: Foreign Trade Forecasting, Explainable Artificial Intelligence (xAI), SHAP Attribution, Ridge Regression, Econometric Feature Engineering, India Trade Volatility.

1. Introduction

International trade dynamics govern national income accounts, exchange rate trajectories, and sovereign balance-of-payments equilibrium. For emerging economies such as India, foreign merchandise trade has undergone structural transformations over the past three decades. Following the 1991 Liberalization, Privatization, and Globalization (LPG) policy reforms, India's trade openness expanded rapidly. However, integration into global value chains has increased exposure to external shocks, including the

2008 Global Financial Crisis (GFC), domestic structural adjustments, and the supply chain disruptions of the 2020 COVID-19 pandemic.

Predicting aggregate trade flows remains challenging due to non-linear dependencies, seasonal production variations, macro-level price transmission delays, and volatility spikes. Traditional econometrics, such as Vector Autoregressions (VAR) and Auto-Regressive Integrated Moving Average (ARIMA) models, provide clear structural interpretations based on economic theory. However, they frequently rely on strict linear assumptions, stationarity transformations, and Gaussian distribution properties that can underfit multi-decade structural shifts.

Conversely, advanced machine learning ensembles (such as Gradient Boosted Trees and Random Forests) capture complex non-linear interactions without requiring rigid parameter distributions. Despite these advantages, machine learning models face two major limitations in time-series macroeconomics:

1. The "Black-Box" Interpretability Deficit: Complex ensembles obscure decision pathways, making it difficult for policy analysts to extract transparent economic insights or isolate structural volatility drivers.
2. Out-of-Sample Extrapolation Failure: Un-differenced tree ensembles struggle to extrapolate deterministic upward macroeconomic trends beyond historical training boundaries.

   To address these challenges, this study develops a hybrid machine learning and xAI interpretability framework applied to India's monthly export and import series spanning 1990 to 2025. By unifying autoregressive lag mechanics with Shapley Additive exPlanations (SHAP), we evaluate predictive accuracy while isolating the key macroeconomic drivers of trade performance.

2. Literature Review

The empirical modeling of international trade flows has evolved from foundational gravity models and linear autoregressive frameworks toward machine learning architectures. Early structural contributions focused on demand-side price and income elasticities, establishing trade balance persistence through autoregressive lag structures. With the expansion of high-frequency financial datasets, researchers increasingly adopted time-series models. Box, Jenkins, Reinsel, and Ljung [1] popularized ARIMA modeling for macroeconomic forecasting, establishing univariate benchmarks. To account for interdependent dynamic relationships between foreign imports and national exports, Sims [2] introduced Vector Autoregressions (VAR), providing a framework for impulse response analysis. However, Stock and Watson [4] noted that standard linear models struggle when structural breaks alter long-term economic relationships, highlighting the need for dynamic specifications.

Recent literature explores machine learning applications in macroeconomic forecasting. White [11] noted that machine learning architectures provide systematic frameworks for capturing non-linear relationships in macroeconomic forecasting. Tree-based ensembles—such as Random Forest introduced by Breiman [7], XGBoost developed by Chen and Guestrin [5], and LightGBM by Ke et al. [6]—can outperform classic linear representations by capturing non-linear interactions without requiring explicit model specification. In empirical trade research, these models have been applied to commodity price predictions, foreign exchange volatility, and regional export demand patterns.

Despite these empirical gains, adoption by central banks and economic policy institutions remains cautious due to transparency concerns. This has driven interest in Explainable AI (xAI). Ribeiro, Singh, and Guestrin [10] emphasized the necessity of model explanations for establishing trust in complex

predictions. Lundberg and Lee [3] unified game-theoretic concepts by introducing SHAP (Shapley Additive exPlanations), which assigns additive feature importance values derived from cooperative game theory. Molnar [9] provided a broader taxonomy of interpretable machine learning methods, categorizing local and global attribution frameworks. For evaluating predictive performance over time, Diebold and Mariano [12] established standard test statistics for comparing out-of-sample forecast accuracy. In linear regularization, Hoerl and Kennard [8] developed Ridge Regression to resolve multicollinearity in highly correlated feature spaces. Applying SHAP interpretability to multi-decade macro-trade series with rolling volatility metrics remains relatively underexplored in empirical trade literature.

3. Methodology

3.1 Dataset Description and Preprocessing

The dataset comprises 840 monthly observation records of India's foreign merchandise trade, divided into 420 monthly aggregate Export figures and 420 Import figures spanning 35 financial years (1990–1991 to 2024–2025). Values are measured in Billions of USD ($B). Financial year notation (April to March) was converted into continuous calendar timestamps. The dataset was restructured into a wide-format monthly time series vector Y_t = [Exports_t, Imports_t]^T for t = 1, 2, ..., 420.

3.2 Feature Engineering Architecture

To model dynamic dependencies, seasonality, and macro shocks, the feature space was expanded across five domain categories:

1. Autoregressive Lags: Lagged terms for exports (Exports_{t-k}) and imports (Imports_{t-k}) where k in {1, 2, 3, 6, 12}.

2. Moving Averages: Rolling expectations over short, medium, and long horizons:

$$MA_{(w, t)} = (1 / w) * sum_{(i=1)}^{w} Exports_{(t-i)}, \quad w \in \{3, 6, 12\} \qquad \text{..... Eq. 1}$$

3. Rolling Volatility: Moving standard deviations capturing short-term variability:

$$Vol_{(w, t)} = sqrt((1 / (w - 1)) * sum_{(i=1)}^{w} (Exports_{(t-i)} - MA_{(w, t)})^2), \quad w \in \{3, 6\} \text{..... Eq. 2}$$

4. Derived Econometric Metrics:

Lagged Trade Balance: TB_{t-1} = Exports_{t-1} - Imports_{t-1}

Lagged Coverage Ratio: CR_{t-1} = Exports_{t-1} / (Imports_{t-1} + epsilon)

Growth Rates: Month-over-Month (MoM_{t-1}) and Year-over-Year (YoY_{t-1}) percentage shifts.

5. Structural Shock Indicators: Binary dummy variables capturing structural policy shifts and global crises:

$$Shock_{LPG} = I(t >= \text{July 1991}) \qquad \text{..... Eq. 3}$$

$$Shock_{GFC} = I(\text{Sept 2008} <= t <= \text{Dec 2009}) \qquad \text{..... Eq. 4}$$

$$Shock_{COVID} = I(\text{March 2020} <= t <= \text{June 2021}) \qquad \text{..... Eq. 5}$$

3.3 Chronological Validation Strategy

To prevent data leakage in time-series forecasting, we avoid random k-fold cross-validation. Instead, we use a chronological out-of-sample test split:

Training Window: April 1990 to March 2020 (360 months; 85.7% of data).

Out-of-Sample Test Window: April 2020 to March 2025 (60 months; 14.3% of data).

3.4 Model Formulation

We evaluate linear regularized models alongside non-linear tree ensembles:

Ridge Regression penalizes large coefficients via an L2 norm penalty:

min_beta sum_(i=1)^N (y_i - beta_0 - sum_(j=1)^p x_ij * beta_j)^2 + alpha * sum_(j=1)^p beta_j^2 ..... Eq. 6

Random Forest & Extra Trees Regressors: Ensemble bagging algorithms that build decision trees on random feature sub-spaces.

XGBoost & LightGBM: Gradient-boosted decision tree architectures optimized for tabular inputs.

3.5 Game-Theoretic SHAP Interpretability

To explain predictions, we apply Shapley Additive exPlanations (SHAP), derived from cooperative game theory. Feature values are evaluated across all possible feature subsets S subseteq F \ {j}. The Shapley value phi_j for feature j is computed as:

phi_j(x) = sum_(S subseteq F \ {j}) [(|S|! * (|F| - |S| - 1)!) / (|F|!)] * [f_x(S union {j}) - f_x(S)]... Eq. 7

For linear models, SHAP values simplify to the feature value deviation from its training mean, weighted by its regularized regression coefficient:

phi_j(x) = beta_j * (x_j - E[x_j]) ..... Eq. 8

4. Results & Discussions

4.1 Comparative Model Performance

Table 1 outlines model accuracy on out-of-sample test data (April 2020 to March 2025).

Table 1: Model Performance Benchmark (Out-of-Sample Test: 2020–2025)

Model	RMSE ($B)	MAE ($B)	MAPE (%)	R^2 Score
Ridge Regression	**3.6592**	**2.6889**	**10.05**	**0.6278**
Extra Trees	7.4762	6.7060	20.67	-0.5538
Random Forest	7.7578	6.9807	21.56	-0.6731
LightGBM	8.2962	7.5281	23.04	-0.9134
XGBoost	8.6527	7.8925	24.16	-1.0813

Ridge Regression achieves the highest predictive accuracy, recording an RMSE of $3.6592B, a MAE of $2.6889B, a MAPE of 10.05%, and an R^2 of 0.6278.

Conversely, non-linear tree ensembles exhibit negative R^2 values. This occurs because decision-tree partitions rely on orthogonal feature splits bounded by historical training thresholds. When applied to non-stationary macroeconomic series with long-term upward trends, un-differenced tree models cannot extrapolate predictions beyond the maximum values observed in the training set (1990–2020). Consequently, they underpredict export expansions during the 2021–2025 post-pandemic growth period. In contrast, Ridge Regression's linear structural form captures deterministic global trends effectively.

4.2 Global SHAP Feature Importance

Because Ridge Regression achieved superior predictive performance, its decision pathways were interpreted using linear SHAP attribution. The global feature importance ranking derived from SHAP analysis highlights key structural insights:

Figure 1 illustrates the global feature importance ranked by mean absolute SHAP values ($B), revealing that long-term moving averages (Exports_MA_12 and Imports_MA_12) and short-term autoregressive momentum serve as the primary drivers of India's aggregate export predictions.

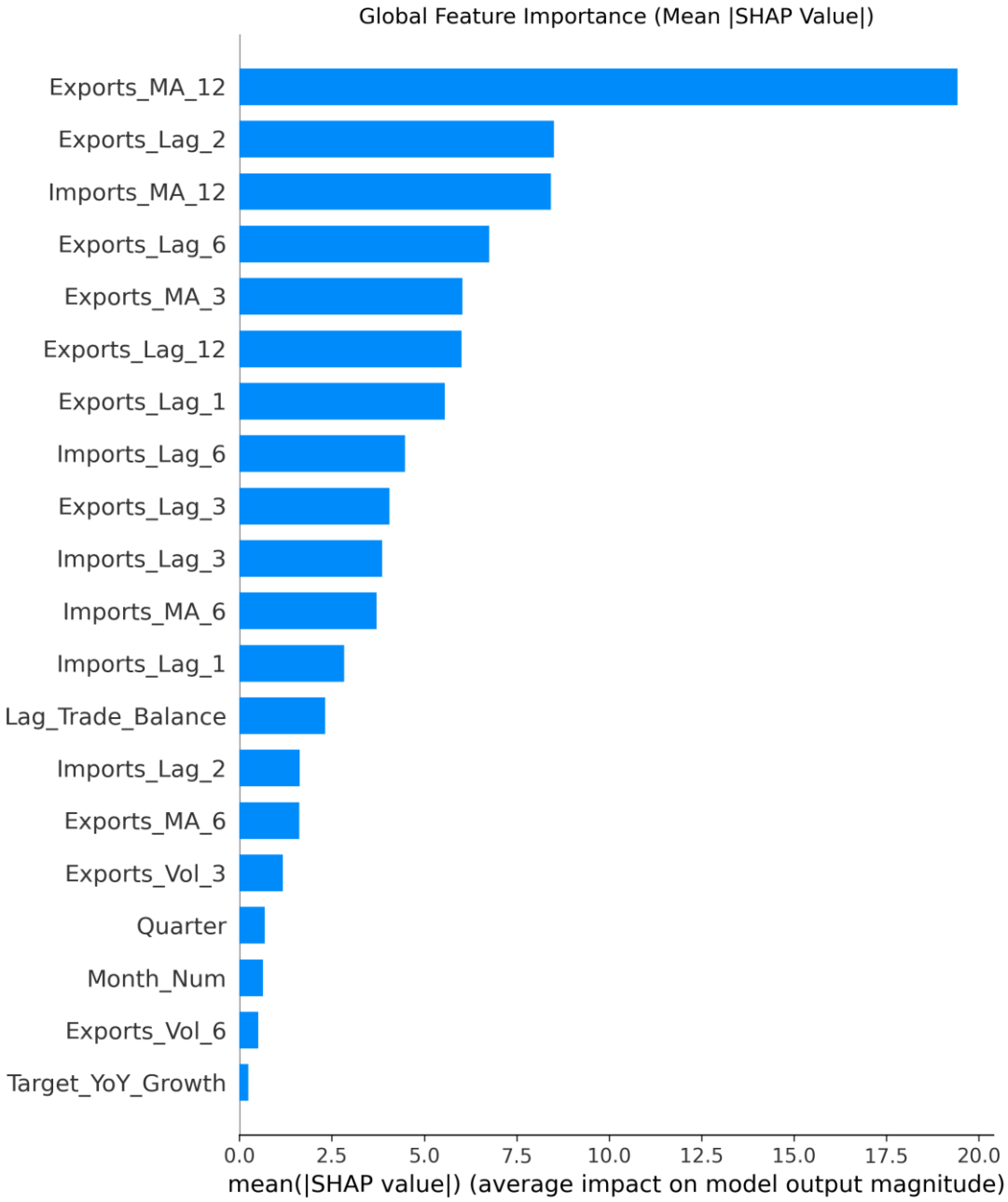


Figure 1: Global Feature Importance based on Mean Absolute SHAP Values ($B).

1. Dominance of Long-Term Moving Expectations: Exports_MA_12 emerges as the primary global driver, contributing a mean absolute SHAP value of $19.42 billion. This indicates that annual baseline expectations establish the underlying trajectory for monthly trade figures.
2. Import-Export Structural Coupling: Imports_MA_12 ranks as the third most important global feature ($8.43 billion). This reflects India's structural import dependence, where imported capital goods, raw materials, and intermediate inputs drive subsequent export manufacturing and assembly.
3. Autoregressive Momentum Lags: Medium and short-term lags (Exports_Lag_2 at $8.54 billion, Exports_Lag_6 at $6.77 billion, and Exports_Lag_1 at $5.56 billion) act as key fine-tuning mechanisms, adjusting monthly forecasts around the annual baseline.

4.3 Directional Volatility Dynamics via SHAP Beeswarm Interpretations

The SHAP Beeswarm analysis illustrates both the direction and magnitude of feature impacts on export predictions:

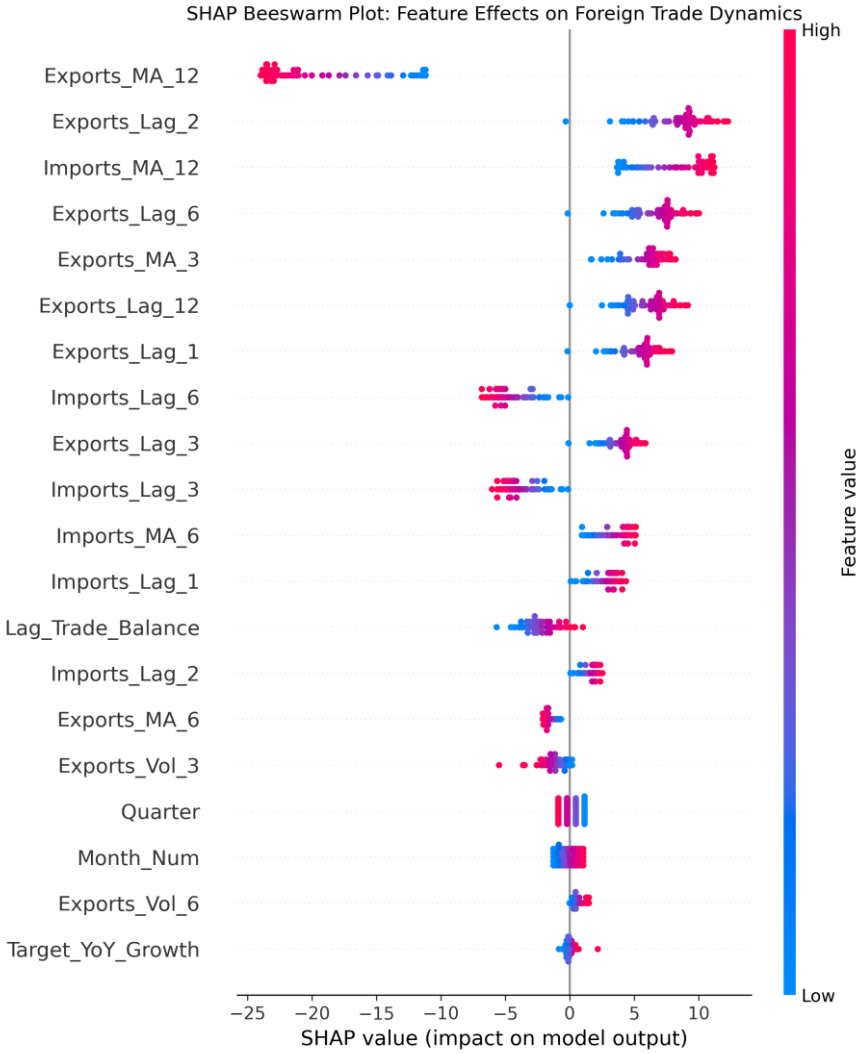


Figure 2: SHAP Beeswarm Plot illustrating directional feature effects on foreign trade dynamics.

Figure 2 presents the SHAP beeswarm plot depicting the directional impact and magnitude of individual feature values, demonstrating that elevated long-term trend baselines drive positive export expansions while spikes in short-term rolling volatility exert downward pressure on trade forecasts.

1. Trend Continuations vs. Mean Reversion: High values of Exports_MA_12 yield positive SHAP values (up to +12 billion USD impact), confirming that long-term structural expansions drive higher export forecasts.
2. Short-Term Volatility Pressures: High short-term rolling volatility (Exports_Vol_3) corresponds to negative SHAP values during contractionary periods. This confirms that spikes in monthly export variance create downward pressure on forecast expectations.
3. Trade Deficit Adjustments: Negative values in Lag_Trade_Balance (representing widening import deficits) shift predictions downward, reflecting the contractionary pressure of trade balance deficits on short-term export momentum.

5. Conclusion

This study provides an empirical framework combining time-series modeling with Explainable Artificial Intelligence (xAI) to analyze India's foreign trade dynamics from 1990 to 2025.
The empirical findings yield three key conclusions:

1. Model Selection in Non-Stationary Series: Regularized linear models like Ridge Regression can outperform non-linear tree ensembles on out-of-sample macroeconomic time series, as un-differenced tree models struggle with trend extrapolation beyond historical boundaries.
2. Structural Trade Dependencies: Game-theoretic SHAP interpretations demonstrate that 12-month moving export and import baselines serve as the primary drivers of global trade performance, while short-term autoregressive lags fine-tune monthly variance.
3. Policy and Risk Management Insights: Integrating SHAP interpretability into econometric models provides transparent attribution for policy analysis, helping identify whether trade fluctuations stem from long-term structural trends or short-term external shocks.

References

1. G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, Time Series Analysis: Forecasting and Control, 5th ed. Hoboken, NJ, USA: Wiley, 2015.
2. C. A. Sims, "Macroeconomics and Reality," Econometrica, vol. 48, no. 1, pp. 1–48, Jan. 1980.
3. S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," in Advances in Neural Information Processing Systems (NeurIPS 30), I. Guyon et al., Eds. Curran Associates, Inc., 2017, pp. 4765–4774.
4. J. H. Stock and M. W. Watson, "Vector Autoregressions," Journal of Economic Perspectives, vol. 15, no. 4, pp. 101–115, Fall 2001.
5. T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (KDD '16), San Francisco, CA, USA, 2016, pp. 785–794.
6. G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "LightGBM: A Highly Efficient Gradient Boosting Decision Tree," in Advances in Neural Information Processing Systems (NeurIPS 30), I. Guyon et al., Eds. Curran Associates, Inc., 2017, pp. 3146–3154.
7. L. Breiman, "Random Forests," Machine Learning, vol. 45, no. 1, pp. 5–32, Oct. 2001.
8. A. E. Hoerl and R. W. Kennard, "Ridge Regression: Biased Estimation for Nonorthogonal Problems," Technometrics, vol. 12, no. 1, pp. 55–67, Feb. 1970.

9. C. Molnar, Interpretable Machine Learning: A Guide for Making Black Box Models Explainable, 2nd ed. Munich, Germany: Leanpub, 2022.
10. M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why Should I Trust You?': Explaining the Predictions of Any Classifier," in Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (KDD '16), San Francisco, CA, USA, 2016, pp. 1135–1144.
11. H. White, "Systematic Economic Forecasting and Machine Learning," Journal of Economic Literature, vol. 44, no. 2, pp. 431–458, Jun. 2006.
12. F. X. Diebold and R. S. Mariano, "Comparing Predictive Accuracy," Journal of Business & Economic Statistics, vol. 13, no. 3, pp. 253–263, Jul. 1995.