Feature Learning Behavior in Trade Forecasting: Evidence from Tree-Based Machine Learning Models

S. Balasubramanian
Department of Computer and Information Science, Annamalai University, Tamil Nadu, India
spbala12@gmail.com (corresponding author)

M. Natarajan
Department of Computer and Information Science, Annamalai University, Tamil Nadu, India
mnid.2004@gmail.com

Received: 17 January 2026 | Revised: 9 February 2026 and 19 February 2026 | Accepted: 22 February 2026

 DOI: https://doi.org/10.48084/etasr.17612

ABSTRACT

This study investigates how tree-based machine learning models learn from historical trade data by examining the importance of lagged export and import values in the context of India's monthly aggregate trade series from Fiscal Year (FY) 1990–91 to FY 2024–25. Rather than prioritizing predictive accuracy, the analysis focuses on feature-based learning behavior using lagged inputs constructed within a supervised learning framework. A time-aware training–testing split is employed to preserve chronological integrity. Feature importance analysis is treated as the primary analytical outcome to evaluate how models allocate learning weights across short-term and medium-term temporal inputs. The results show that recent export lags consistently dominate model learning, whereas selected medium-term import lags also contribute meaningfully. Robustness checks using alternative lag specifications confirm the stability of short-term dominance. These findings demonstrate that tree-based trade forecasting models learn in a structured and selective manner rather than distributing attention uniformly across historical observations. By emphasizing interpretability over pure accuracy comparisons, the study provides deeper insight into temporal learning dynamics in machine learning-based trade forecasting.

Keywords-trade forecasting; feature learning; lagged variables; tree-based machine learning; XGBoost; model interpretability

I. INTRODUCTION

International trade forecasting plays a central role in economic planning, policy formulation, and strategic decision-making. Export and import flows exhibit strong temporal dependence, where past trade values influence future outcomes. For this reason, forecasting models typically rely on historical observations to capture the dynamics embedded in trade series. With the increasing availability of structured economic data, machine learning approaches have gained prominence due to their flexibility and ability to model nonlinear relationships.

Among these methods, tree-based ensemble techniques have become widely adopted. Extreme Gradient Boosting (XGBoost), introduced by authors in [1], provides a scalable and efficient tree-boosting framework capable of capturing complex nonlinear interactions. The theoretical foundation of ensemble tree learning was established by authors in [2] through the Random Forest algorithm, which demonstrated that tree ensembles learn selectively from input variables rather than treating all inputs equally. As model complexity increased, interpretability emerged as a critical concern. Shapley Additive Explanations (SHAP)-based methods provide a unified framework for interpreting model predictions [3], whereas TreeSHAP enables consistent feature attribution specifically for tree ensembles [4]. Permutation-based importance measures further contribute to assessing variable relevance in predictive models [5].

In the broader forecasting domain, the M4 competition highlighted the growing role of machine learning methods and emphasized the need for methodological rigor in forecasting applications [6]. Several studies have applied machine learning techniques to international trade forecasting, confirming their empirical relevance while focusing primarily on predictive performance [7, 8]. More recent work has extended these approaches to real-time trade monitoring and nowcasting contexts [9]. Related research published in this journal has also demonstrated the applicability of machine learning models to trade and economic analysis [10]. Authors in [11] provide a comprehensive discussion of interpretability methods, emphasizing the importance of understanding internal model behavior rather than relying solely on accuracy metrics.

Despite the expanding literature, limited attention has been devoted to examining how tree-based models internally learn from lagged temporal inputs in trade forecasting settings. Lagged export and import variables are commonly introduced to improve predictive accuracy, yet their relative contribution to model learning is rarely analyzed in a structured manner. While ensemble methods are known to allocate importance selectively across inputs, the distribution of learning weights across short-term and medium-term trade memory remains insufficiently explored.

This study addresses this gap by investigating feature learning behavior in tree-based models applied to India's monthly aggregate export and import data from Fiscal Year (FY) 1990–91 to FY 2024–25. Instead of prioritizing forecasting accuracy, feature importance is treated as the primary analytical outcome to evaluate how learning weights are distributed across lagged inputs. The analysis focuses on identifying whether short-term lags dominate model learning, whether import and export lags exhibit distinct importance structures, and whether these patterns remain stable under alternative lag specifications.

The study contributes in three ways. First, it shifts the analytical emphasis in trade forecasting research from predictive performance comparison toward internal learning structure in machine learning models. Second, it employs feature importance as a central evaluative framework to strengthen interpretability within applied forecasting. Third, it provides empirical evidence on how temporal trade memory is prioritized within a tree-based learning system in the Indian trade context.

The objectives of the study are to examine how tree-based machine learning models learn from historical trade data, to analyze the relative importance of short-term and medium-term lagged inputs, and to assess the stability of importance rankings under alternative lag configurations. Based on these objectives, the following hypotheses are formulated:

- H1: Short-term lagged trade values exert greater influence on model learning than longer-term lagged values.
- H2: Import lags exhibit distinct temporal importance patterns relative to export lags.
- H3: The dominance of short-term lags remains stable across alternative lag specifications.

II. DATA DESCRIPTION

The study uses monthly aggregate export and import data for India, covering the period from FY 1990–91 to FY 2024–25. The dataset consists of 420 monthly observations, with trade values reported in USD billion. The data were obtained from Indiastat.com, which compiles official trade statistics from national sources such as the Ministry of Commerce and Industry, Government of India. The dataset represents continuous monthly observations of total exports and imports and was examined for missing values and inconsistencies prior to analysis. Basic preprocessing steps were applied to ensure temporal continuity and consistency across the series, making the dataset suitable for lag-based time series modeling.

The dataset comprises two primary time series: total monthly exports and total monthly imports. Each observation corresponds to a single calendar month, allowing the data to be treated as a structured time series suitable for lag-based modeling. Prior to analysis, the series were examined for missing values and inconsistencies. No external macroeconomic variables were incorporated, as the objective of the study is to isolate learning behavior based solely on historical trade information.

To enable machine learning implementation, the original time series were transformed into a supervised learning format through lag construction. Past export and import values were used as explanatory variables, whereas the trade value in month t was defined as the prediction target. This transformation allows the model to learn temporal dependencies directly from historical observations.

A chronological training–testing split was applied to preserve temporal integrity and prevent information leakage. The first 80% of observations (336 months) were used for model training, whereas the remaining 20% (84 months) were reserved for out-of-sample evaluation. This time-aware splitting strategy reflects realistic forecasting conditions and supports methodological transparency.

The processed monthly export and import dataset used in this study has been made publicly available through an open-access repository to enhance transparency and reproducibility [12]. The key variables used in the analysis are summarized in Table I.

TABLE I. DESCRIPTION OF TRADE DATA VARIABLES

Variable	Description	Unit	Role
Export_t	Monthly aggregate exports (India)	USD billion	Target / lagged input
Import_t	Monthly aggregate imports (India)	USD billion	Lagged input

III. FEATURE ENGINEERING AND LAG CONSTRUCTION

Feature engineering is designed to capture the temporal dependence inherent in monthly trade flows. Since current export and import values are influenced by prior observations, historical trade values are structured as lagged inputs within a supervised learning framework.

For each month t, lagged values from previous months are constructed for both export and import series. In the baseline specification, six monthly lags are included for each variable. This lag horizon is selected to represent short-term persistence (1–3 months) and medium-term trade adjustments (4–6 months), reflecting production cycles, shipment delays, and inventory effects commonly observed in aggregate trade systems.

Lagged features are created by shifting the original series backward in time. The lagging process introduces missing values at the beginning of the sample, and these initial observations are removed to maintain a consistent training dataset. The chronological order of observations is strictly preserved.

No additional scaling, transformation, or derived indicators are introduced. This restriction is intentional and aligns with the study's objective of isolating temporal learning behavior based solely on observed trade values. By keeping each feature directly linked to a specific past month, the resulting feature importance measures can be interpreted transparently in terms of historical trade memory.

To evaluate the robustness of the baseline lag specification, alternative lag structures using three and nine monthly lags are also examined. This comparison allows assessment of whether the dominance of short-term lags remains stable across different temporal configurations.

IV. METHODOLOGY AND EXPLAINABILITY FRAMEWORK

The modeling framework employs XGBoost regression [1] to examine how tree-based ensemble methods learn from lagged trade inputs. The objective is not solely to minimize forecasting error, but to analyze how learning weights are allocated across temporal features.

A. Model Specification

The forecasting task is formulated as a supervised regression problem. Let y_t denote the trade value at time t, and let x_{t-i} represent the lagged trade value at lag i. The learning structure is expressed as:

$$y_t = f(x_{t-1}, x_{t-2}, \dots, x_{t-k}) + \varepsilon_t \quad (1)$$

where $f(\cdot)$ denotes the nonlinear function estimated by the XGBoost ensemble, k is the maximum lag length, and ε_t represents the residual term. The function $f(\cdot)$ is constructed through an additive sequence of decision trees optimized using gradient boosting.

B. Training Procedure

The model is trained using a chronological split to preserve temporal order. The first 336 observations (80%) are used for training and the remaining 84 observations (20%) are reserved for out-of-sample evaluation. This time-aware approach prevents information leakage and reflects realistic forecasting conditions.

C. Feature Importance Measurement

To interpret model learning behavior, feature importance is evaluated using the gain metric provided by XGBoost. Gain measures the average reduction in the objective function attributable to splits on a given feature across all trees in the ensemble. A higher gain value indicates that the feature contributes more substantially to improving model fit.

Feature importance scores are extracted for each lagged export and import variable. By examining how gain values are distributed across lag lengths, the analysis reveals how the model prioritizes short-term versus medium-term trade memory.

D. Stability and Interpretability

To assess robustness, feature importance rankings are examined across repeated model runs and alternative lag configurations. Consistent importance patterns indicate stable learning behavior, whereas significant variation would suggest sensitivity to lag specification. Because each feature corresponds directly to an observed trade value from a specific past month, importance scores can be interpreted transparently in terms of temporal trade memory.

The modeling design intentionally restricts inputs to historical trade values without incorporating external macroeconomic variables. This controlled structure isolates endogenous temporal learning patterns and strengthens interpretability.

V. RESULTS AND DISCUSSION

The XGBoost model was trained using six monthly lags for both export and import series. Feature importance was evaluated using the gain metric, which measures the average reduction in the objective function contributed by splits on each feature across all trees in the ensemble.

Table II reports the feature importance values ranked in descending order. The distribution of importance is highly uneven, indicating selective allocation of learning weights across temporal inputs.

TABLE II. FEATURE IMPORTANCE RANKING (GAIN METRIC)

Rank	Feature	Importance
1	Export_lag_1	0.321709
2	Import_lag_4	0.283721
3	Import_lag_5	0.192506
4	Export_lag_5	0.057294
5	Import_lag_6	0.031814
6	Export_lag_2	0.030502
7	Import_lag_1	0.025065
8	Import_lag_2	0.024759
9	Import_lag_3	0.018878
10	Export_lag_6	0.005465
11	Export_lag_3	0.004445
12	Export_lag_4	0.003844

Export_lag_1 exhibits the highest importance score, indicating that the most recent export value plays a dominant role in guiding model learning. In contrast, higher-order export lags contribute only marginally. Import-related lags, however, display a distinct temporal pattern, with Import_lag_4 and Import_lag_5 ranking immediately after Export_lag_1. This ordering indicates that import activity influences learning with a delay rather than through immediate contemporaneous effects.

Figure 1 visualizes the same importance distribution. A sharp decline in importance is observed as lag length increases, confirming that the model concentrates learning weights on a limited subset of recent and strategically positioned medium-term lags.

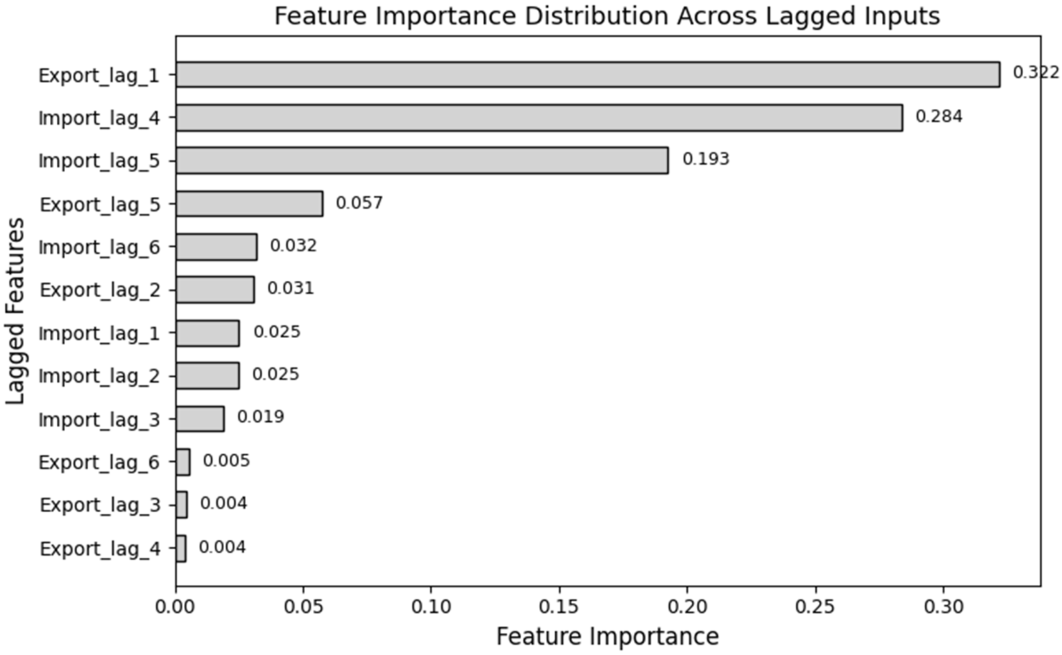

Fig. 1. Feature importance distribution across lagged inputs.

From an economic perspective, the dominance of Export_lag_1 reflects strong short-term persistence and momentum effects in monthly trade flows, often driven by shipment cycles and reporting continuity. The prominence of medium-term import lags may capture delayed transmission mechanisms, where prior import activity affects subsequent export performance through production inputs, inventory adjustments, and supply-chain linkages.

The results extend beyond confirming temporal dependence. The key insight lies in the structured allocation of learning weights. The model does not distribute importance uniformly across all historical observations; instead, it selectively prioritizes recent export values and specific medium-term import lags, internalizing trade memory within nonlinear decision structures.

A. Robustness Analysis

To assess stability, alternative lag specifications using three and nine monthly lags were examined. Across these configurations, short-term lags consistently retained dominant importance rankings, whereas higher-order lags remained marginal. Although absolute importance values varied slightly, the relative hierarchy of temporal importance remained stable. This robustness supports Hypothesis H3 and indicates that short-term dominance is not driven by a specific lag configuration.

B. Implications

Feature importance analysis provides interpretative value beyond predictive accuracy metrics. By identifying which historical observations guide model learning, the approach enhances transparency in machine learning-based trade forecasting. The findings also suggest that expanding lag length indiscriminately may increase model complexity without substantially altering the underlying learning structure.

VI. SCOPE AND DESIGN BOUNDARIES

The study is intentionally designed to isolate temporal learning behavior in tree-based models using historical trade values as the sole inputs. External macroeconomic variables such as exchange rates, inflation, or global demand indicators are not incorporated. This restriction is deliberate and allows clear attribution of feature importance to endogenous trade memory rather than to broader economic factors.

The lag structure is fixed within a predefined monthly horizon to provide a controlled environment for examining short-term and medium-term trade persistence. While alternative lag lengths are explored for robustness, the analysis does not extend to adaptive or high-dimensional lag selection frameworks, as the primary objective is interpretability rather than model complexity.

The modeling framework focuses on tree-based ensemble methods due to their transparency in feature importance extraction. The findings therefore reflect learning behavior within this class of models and are not intended as universal claims across all algorithmic families. Similarly, feature importance is interpreted as a measure of contribution to model fit rather than as evidence of economic causality.

Finally, the analysis is conducted at the aggregate trade level. Disaggregated sectoral or bilateral trade structures may exhibit distinct temporal dynamics, which represent a potential avenue for further investigation.

Within these defined boundaries, the study provides a focused and interpretable examination of how machine learning models internalize trade memory in a structured forecasting setting.

VII. CONCLUSION AND FUTURE SCOPE

This study investigated how tree-based machine learning models internalize temporal trade information by examining the importance of lagged export and import values in India's monthly trade series from Fiscal Year (FY) 1990–91 to FY 2024–25. Rather than emphasizing forecasting accuracy, the analysis focused on feature importance as a means of understanding internal learning behavior.

The findings indicate that model learning is highly selective. Short-term export lags dominate the learning structure, whereas specific medium-term import lags contribute meaningfully. Higher-order lags exhibit limited influence. These results suggest that tree-based models prioritize recent trade memory while incorporating delayed import effects in a structured manner. The contribution of the study lies not in confirming temporal dependence, but in demonstrating how such dependence is operationalized within nonlinear ensemble learning frameworks.

By treating feature importance as the primary analytical outcome, the study advances interpretability in machine learning-based trade forecasting. The results illustrate how explainability tools can clarify internal model structure and provide insight beyond conventional accuracy metrics. This perspective contributes to ongoing discussions on transparency and accountability in applied machine learning for economic analysis.

Future research may extend this framework by examining whether similar temporal learning hierarchies emerge across different countries, trade regimes, or levels of disaggregation. Investigating how endogenous trade memory interacts with external macroeconomic shocks could provide deeper theoretical insight into model-based representations of trade dynamics. Comparative analysis across alternative algorithmic families may also clarify whether structured temporal prioritization is specific to tree-based ensembles or reflects broader machine learning behavior in economic forecasting contexts.

REFERENCES

[1] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, CA, USA, 2016, pp. 785–794, https://doi.org/10.1145/2939672.2939785.

[2] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct. 2001, https://doi.org/10.1023/A:1010933404324.

[3] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, CA, USA, 2017, pp. 4768–4777, https://doi.org/10.48550/arXiv.1705.07874.

[4] S. M. Lundberg, G. G. Erion, and S.-I. Lee, "Consistent Individualized Feature Attribution for Tree Ensembles." arXiv, Mar. 07, 2019, https://doi.org/10.48550/arXiv.1802.03888.

[5] A. Altmann, L. Toloşi, O. Sander, and T. Lengauer, "Permutation importance: a corrected feature importance measure," *Bioinformatics*, vol. 26, no. 10, pp. 1340–1347, May 2010, https://doi.org/10.1093/bioinformatics/btq134.

[6] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "The M4 Competition: 100,000 time series and 61 forecasting methods," *International Journal of Forecasting*, vol. 36, no. 1, pp. 54–74, Jan. 2020, https://doi.org/10.1016/j.ijforecast.2019.04.014.

[7] F. Batarseh, M. Gopinath, G. Nalluru, and J. Beckman, "Application of Machine Learning in Forecasting International Trade Trends." arXiv, Oct. 07, 2019, https://doi.org/10.48550/arXiv.1910.03112.

[8] H. Jošić and B. Žmuk, "A Machine Learning Approach to Forecast International Trade: The Case of Croatia," *Business Systems Research Journal*, vol. 13, no. 3, pp. 144–160, Dec. 2022, https://doi.org/10.2478/bsrj-2022-0030.

[9] M. D. Chinn, B. Meunier, and S. Stumpner, "Nowcasting World Trade with Machine Learning: a Three-Step Approach." National Bureau of Economic Research, June 2023, https://doi.org/10.3386/w31419.

[10] J. Sun, Y. Suo, S. Park, T. Xu, Y. Liu, and W. Wang, "Analysis of Bilateral Trade Flow and Machine Learning Algorithms for GDP Forecasting," *Engineering, Technology & Applied Science Research*, vol. 8, no. 5, pp. 3432–3438, Oct. 2018, https://doi.org/10.48084/etasr.2311.

[11] C. Molnar, *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*, 3rd ed. Berlin, Germany: Self-Published, 2022.

[12] B. Subramanian, "Monthly Aggregate Export and Import Trade Data for India (FY 1990–91 to FY 2024–25)." Zenodo, Feb. 19, 2026, https://doi.org/10.5281/zenodo.18699401.