CARMA 2026

8th Int. Conf. on Advanced Research Methods and Analytics
1-3 July 2026, Valencia, Spain
DOI: https://doi.org/10.4995/CARMA2026.2026.21864

GDP Components Nowcasting with Machine Learning

Shereena Silva Salas , **Dennis Malabuyoc Bautista**
Department of Economic Research, Economic and Financial Forecasting Group, Bangko Sentral ng Pilipinas, Philippines.

How to cite: Silva Salas, S.; Malabuyoc Bautista, D. (2026). GDP Components Nowcasting with Machine Learning. In: 8th International Conference on Advanced Research Methods and Analytics (CARMA 2026). Valencia, 1-3 July 2026. https://doi.org/10.4995/CARMA2026.2026.21864

Abstract

Using Philippine data, this study develops a unified machine learning (ML) nowcasting framework for GDP expenditure components. Five ML algorithms (i.e., LASSO, Elastic Net, Random Forest, XGBoost, and Support Vector Regression) are applied and ensemble models are constructed. To address the interpretability of ML predictions, SHapley Additive exPlanations (SHAP) are applied to identify the key drivers of nowcasts for each component. Results demonstrate the potential of ML methods relative to benchmarks, with component-specific patterns in predictive accuracy and feature importance aligned with economic theory. The framework offers a practical tool for real-time macroeconomic surveillance and policy analysis.

Keywords: *Nowcasting; Machine Learning; GDP Components; Explainable AI*

1. Introduction

Accurate, timely assessments of current economic conditions are essential for effective monetary and fiscal policy, yet official GDP releases are published with significant lags. Nowcasting using mixed-frequency and high-frequency indicators has become a standard tool in central banks to bridge this information gap. In practice, nowcasting frameworks combine a broad set of indicators with different frequencies with statistical and more recently, machine learning (ML) models to produce early estimates of quarterly aggregates before official releases become available.

For an emerging economy such as the Philippines, the value of reliable nowcasts is particularly high because policy decisions must often be taken in the presence of large shocks and structural change. Traditional approaches remain useful but face limitations in high-dimensional and unstable environments.

This paper develops a unified ML-based nowcasting framework for Philippine GDP and its expenditure components: household final consumption expenditure (HFCE), gross fixed capital

Article in Press

This work is licensed under a Creative Commons License CC BY-NC-SA 4.0
Editorial Universitat Politècnica de València

formation (GFCF), government final consumption expenditure (GFCE), exports, and imports. A common pipeline is implemented across components, combining five ML algorithms (i.e., Least Absolute Shrinkage and Selection Operator (LASSO), Elastic Net, Random Forest, extreme gradient boosting (XGBoost), and Support Vector Regression (SVR)) with time-series cross-validation, hyperparameter optimization, and ensembles. To address concerns about interpretability, the framework uses SHapley Additive exPlanations (SHAP) to identify the key drivers of nowcasts at both global and instance level.

The contribution is threefold. First, the paper provides an ML-based nowcasting system that covers all major expenditure-side GDP components for the Philippines within a single, consistent setup. Second, it evaluates the relative performance of regularization-based, tree-based, and kernel-based ML models and ensembles against simple autoregressive benchmarks, highlighting component-specific gains. Third, it demonstrates how SHAP can be used to translate ML nowcasts into economically interpretable narratives.

2. Literature Review

2.1. Nowcasting and High-Dimensional Macroeconomic Models

Nowcasting addresses the policy need for timely GDP assessment despite mixed-frequency indicators, ragged-edge release calendars, and data revisions. In high-dimensional settings—where many predictors compete with short samples—dimension reduction and regularization are commonly used to control overfitting while retaining economically meaningful signals, motivating model-rich forecasting systems that can be updated as new information arrives.

2.2. Machine Learning for Macroeconomic Forecasting and Nowcasting

Evidence increasingly shows that ML methods can outperform traditional time-series benchmarks in macroeconomic forecasting and nowcasting, especially with many heterogeneous predictors and potential nonlinearities. Tree-based models such as Random Forests (Breiman, 2001) and gradient-boosted trees (Chen & Guestrin, 2016), as well as kernel methods such as SVR (Drucker et al., 1996), have delivered gains in high-dimensional and unstable environments (Smeekes & Wijler, 2018; Coulombe et al., 2022) and in GDP nowcasting applications across countries (Bolhuis & Rayner, 2020; Arro-Cannarsa & Scheufele, 2024).

Three themes emerge: performance varies by target and information set; ML is most valuable when relationships are nonlinear or subject to regime change; and forecast combinations often improve robustness. Accordingly, this paper adopts simple ensemble schemes alongside individual model comparisons (Timmermann, 2006).

Article in Press

For the Philippines, Reyes (2022) finds that penalized regression and tree-based models improve liquidity nowcasts relative to standard time-series benchmarks, suggesting scope for ML gains in policy monitoring applications.

This study extends this line of work by implementing a unified ML nowcasting pipeline for Philippine GDP expenditure components, combining systematic model comparison, simple ensembles, and SHAP-based explanations (Lundberg & Lee, 2017) in a real-time evaluation setting.

3. Data and Forecast Setup

The dataset combines quarterly series from the BSP, Philippine Statistics Authority (PSA), and CEIC from 2000Q1 to 2025Q4, yielding about 100 observations and 39 to 78 selected features per component. All target variables are expressed in year-on-year growth rates to reduce seasonality and align the nowcasts with policy-relevant growth concepts used in real-time monitoring. After standard real-time data preprocessing (i.e., lag alignment, missing-value treatment, temporal aggregation, and feature engineering with lag creation of eight lags) the number of predictors differs by GDP component as shown in Table 1. Predictors cover key macroeconomic, financial, real-sector, and external indicators at monthly and quarterly frequency. Model performance is evaluated using rolling-window time-series cross-validation with 60 periods each (15 years) for eight folds; for each model, hyperparameters are tuned with Optuna over 100 trials using cross-validated loss as the objective. Forecast accuracy is assessed using the Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and the Normalized RMSE (NRMSE). While RMSE measures errors in the units of the target variable, NRMSE scales RMSE by the standard deviation of the test set, allowing comparability across components with different volatilities. Out-of-sample performance is reported for 2025Q2 to 2025Q4.

Table 1. Components and number of features.

Component	Selected Features	Feature Engineering
Household Consumption	51	444
Investments	78	660
Government Expenditure	39	354
Exports	57	399
Imports	59	417
GDP	72	597

Article in Press

4. Methodology

The empirical strategy compares traditional time-series benchmarks and ML models for each GDP component. Random Walk, AR(1), AR(4), and ARIMA models serve as benchmarks. We implement regularization-based models — LASSO (Tibshirani, 1996) and Elastic Net (Zou & Hastie, 2005) — which perform automatic variable selection and shrinkage suited to high-dimensional predictor spaces. Tree-based models include Random Forest (Breiman, 2001), which leverages bootstrap aggregation over deep regression trees, and XGBoost (Chen & Guestrin, 2016), which applies gradient boosting with regularization. SVR (Drucker et. al., 1996) rounds out the model set as a kernel-based alternative that is robust in small-sample, high-dimensional environments. All models are implemented using a common pipeline, with hyperparameters tuned via Optuna over 100 trials using time-series cross-validated loss as the objective. Following Bergmeir & Benítez (2012), we apply a rolling-window cross-validation scheme that respects the temporal ordering of observations and prevents look-ahead bias; the initial training window spans 2004Q1 to 2018Q4, expanding by three quarters at each step with a three-quarter-ahead forecast horizon.

To exploit model diversity, three simple ensemble schemes combine individual ML predictions: (i) an equal-weight average of all models, (ii) an inverse-RMSE ensemble that assigns higher weights to models with lower cross-validated RMSE, and (iii) an inverse-rank ensemble based on performance rankings. Forecast combination has been shown to consistently reduce prediction error relative to individual models across a wide range of settings (Timmermann, 2006), providing a strong motivation for adopting ensemble strategies alongside individual models. For interpretability, SHAP (Lundberg & Lee, 2017) is applied to the best-performing models to decompose each nowcast into feature-level contributions, balancing predictive performance with interpretability.

5. Results

5.1. Out-of-Sample Performance Across Components

Figure 1 reports out-of-sample performance for the nowcasting models by GDP expenditure component across RMSE, MAE, NRMSE and MAPE. The results confirm that no single model dominates across all components, underscoring the need for a component-specific nowcasting strategy. Results for NRMSE and MAPE broadly corroborate the RMSE and MAE rankings discussed below.

Ensemble and tree-based models best track private consumption, with inverse rank posting the lowest RMSE and Random Forest the best MAE, while regularization-based models perform poorly. For investment, LASSO and Elastic Net clearly outperform all alternatives, suggesting that penalized regression is particularly effective at exploiting financial and external conditions

Article in Press

captured in the predictor set. Government spending remains the most difficult component to nowcast, with even the best-performing models — AR(1) and Random Forest — posting errors considerably larger than those for other components, reflecting GFCE's policy-driven, exogenous nature.

For exports, AR(4) posts the lowest errors, though penalized regression methods also perform competitively, suggesting that own-lag persistence dominates but regularization can still extract useful signals from global demand and price conditions. Imports represent a clear win for ML, with SVR and ensemble methods substantially outperforming AR benchmarks and random walk, pointing to highly nonlinear import dynamics. Aggregate GDP is best nowcasted by XGB and Random Forest, with regularization-based methods performing the worst, reinforcing the advantage of nonlinear methods for the aggregate series.

Overall, regularization models dominate GFCF, nonlinear and ensemble methods are most valuable for HFCE, imports, and GDP, while simpler benchmarks remain hard to beat for exports and GFCE. Even where predictive improvements are modest, ML methods remain valuable when paired with SHAP, as they facilitate the translation of nowcasts into economically interpretable narratives.

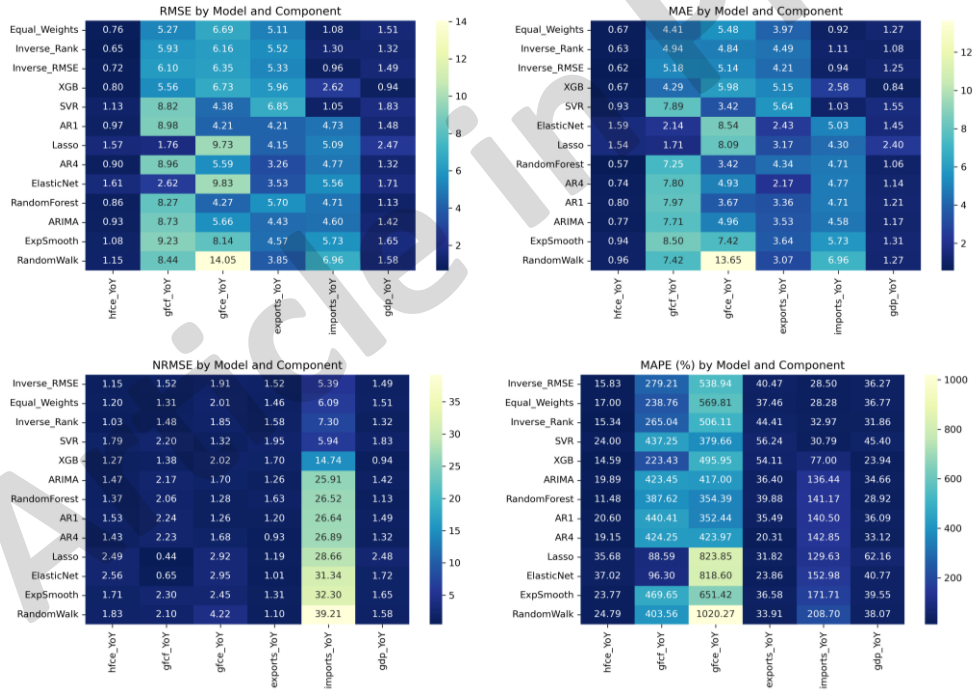


Figure 1. Out-of-sample RMSE and NRMSE per component and per model

Article in Press

5.2 Example component-specific feature importance findings

Focusing on GFCF illustrates how SHAP adds economic structure to ML nowcasts. For the equal-weight ensemble, the beeswarm plot indicates that external demand (major trading partners' GDP), interest rates, and uncertainty measures are among the most influential predictors: higher partner GDP and lower uncertainty tend to push the nowcast up, while higher rates tend to pull it down.

These patterns are consistent with investment dynamics shaped by foreign demand and financing conditions, alongside heightened sensitivity to global and trade policy uncertainty. As with any post-hoc method in a high-dimensional, correlated setting, SHAP attributions are descriptive rather than causal, and feature rankings may vary across estimation windows.

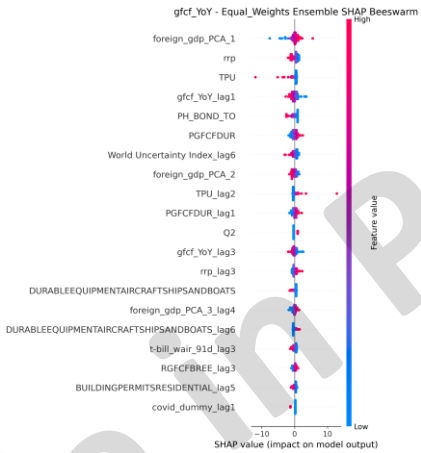


Figure 2. Beeswarm Plots of Shapley Values of GFCF

5.3 Instance-level nowcast decompositions using SHAP

While global SHAP summaries highlight typical drivers, the waterfall plots reveal how these factors combine at specific points in time. Figure 3 presents SHAP waterfall plots for same ensemble nowcasts of GFCF growth from 2025Q4 to 2026Q3, illustrating how quarter-specific information shifts each prediction away from the baseline expectation. In 2026Q1, the nowcast is pulled below the baseline mainly by different policy uncertainty indices and bond turnover ratio. These negative contributions are only partly offset by past bond turnover ratio, high prices for durable equipment, and policy uncertainty from earlier periods which may signal some rebound effect for current investments. In 2026Q2, largely the same factors continue to exert downward pressure, keeping the nowcasts below its baseline but policy uncertainty and past bond turnover ratio some positive support. By 2026Q3, the high policy uncertainty indices from two periods ago emerges as the dominant force pushing the nowcast higher, likely from the

Article in Press

possible rebound effects, outweighing otherwise negative signals from the more recent trade policy uncertainty.

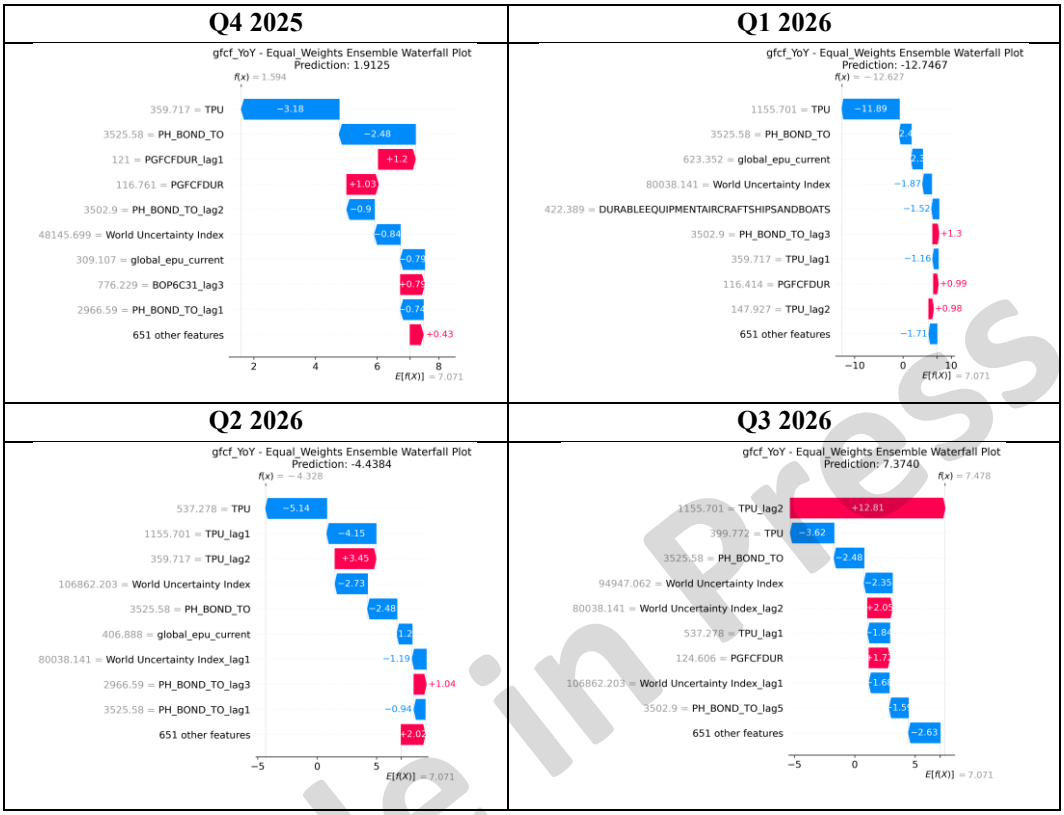


Figure 3. Bar plots and Bee Swarm Plots of Shapley Values of GFCF

6. Conclusion

This paper develops a unified machine-learning-based framework for nowcasting Philippine GDP and its expenditure-side components, combining regularization-based, tree-based, and kernel-based models with ensemble techniques and SHAP-based interpretability. Using a real-time forecasting setup, the analysis shows that no single model dominates across components. Instead, the value of ML is highly component-specific: ML methods and ensemble approaches tend to outperform traditional benchmarks for consumption, investments, imports, and aggregate GDP while offering limited gains for components such as exports and government final consumption expenditure.

Beyond predictive performance, the framework illustrates how interpretability tools can be integrated into ML nowcasting to support economic analysis and real-time monitoring. By

Article in Press

linking component-level nowcasts to economically meaningful drivers, SHAP-based diagnostics help translate high-dimensional ML predictions into policy-relevant narratives, supporting both early detection of turning points and real-time assessment of underlying factors.

The findings suggest that machine-learning methods should be viewed not as a replacement for traditional nowcasting tools, but as a complementary approach. Their value is greatest when combined with rich information set and interpretability methods that help link short-term movements to economically meaningful drivers and policy-relevant insights.

Several extensions remain for future work. Improved filtering of information sets based on economic theory for all components, as well as the incorporation of higher-frequency indicators and alternative data sources (e.g., credit card transactions and domestic policy uncertainty measure), and explicit real-time data vintages (preliminary estimates instead of revised data) may help further improve the performance of ML-based nowcasting frameworks.

References

Arro-Cannarsa, M., & Scheufele, R. (2024). Nowcasting GDP: what are the gains from machine learning algorithms?. *SNB Working Papers*, 06/2024.

Bergmeir, C., & Benítez, J. M. (2012). On the use of cross-validation for time series predictor evaluation. Information Sciences, 191, 192–213.

Bolhuis, M. A., & Rayner, B. (2020). Deus ex machina? A framework for macro forecasting with machine learning. *IMF Working Paper*, WP/20/45.

Breiman, L. (2001). Random forests. Machine Learning, 45, 5–32.

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of KDD 2016, 785–794.

Coulombe P. G., Leroux, M., Stevanovic, D., & Surprenant, S. (2022). How is machine learning useful for macroeconomic forecasting? *Journal of Applied Econometrics*, 37(5), 920–964. https://doi.org/10.1002/jae.2910

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30.

Reyes, J. R. M. (2022). Nowcasting liquidity using machine learning. *BSP Working Paper Series*, 2022-04.

Smeekes, S., & Wijler, E. (2018). Macroeconomic forecasting using penalized regression methods. *International Journal of Forecasting*, 34(3), 408-430. https://doi.org/10.1016/j.ijforecast.2018.01.001

Tibshirani, R. (1996). Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society*: Series B, 58(1), 267–288.

Timmermann, A. (2006). Forecast combinations. *Handbook of Economic Forecasting*, 1, 135–196.

Article in Press

Drucker, H., Burges, C. J., Kaufman, L., Smola, A. J., & Vapnik, V. (1996). Support vector regression machines. In M. C. Mozer, J. D. Jordan, & T. Petsche (Eds.), *Advances in neural information processing systems* 9 (pp. 155–161). MIT Press.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society*: Series B, 67(2), 301–320.

Article in Press