NOWCASTING GDP WITH MACHINE LEARNING: THE CASE OF INDONESIA

GINANJAR UTAMA
NADIRA FIRINDA

2023
MAR

Contents

In this paper, we develop economic models utilizing the concept of Machine Learning to nowcast Indonesia Gross Domestic Product (GDP) growth. The models use various monthly data indicators from 2010 to 2022 that have the potential to be associated with GDP in Indonesia. Thus, the data will be assessed further in terms of their closest relation with the GDP by using six (6) types of association/relationship measure, such as Maximal Information Coefficient (MIC), Maximal Normalized Mutual Information (Max NMI), Kendall, Spearman, Pearson and Distance Correlation, and clustered using unsupervised learning method. The resulting closest variables to describe GDP will be used on four different methods of Machine Learning, which are Elastic Net, Random Forest, XGBoost and Support Vector Machine (SVM) to find the best combined relationship and determine the next quarter of GDP growth, in our case it is the Q4 2022 GDP growth. Our result shows that a range of 4.81% and 5.29% is projected to be the 2022 Q4 GDP growth with the ensemble average of 4.98%. The resulting RMSE varies between 0.03 and 0.49 using the first month of 2010 until the 12th month of 2021 as the training data. While for the testing data, which starts at January up to September of 2022, the RMSE is between 0.25 and 1.36. Additionally to enhance the understanding of the model projection, especially to make the projection easier to be interpreted and explained, we also made the decomposition and the Shapley value of our projection result. Furthermore, we also found that some indicators, namely Current Incomes Index, Consumer Confidence Index, Purchase of Durable Goods Index, PMI, and real retail sales indicators such as: Information and Communication equipment, Food and Beverages, Vehicle Parts, Clothing and Other Goods, have constantly high contributions value in regards to GDP of Indonesia. The results show that this model can accurately predict GDP growth in both normal and pandemic periods.

Introduction

Economic projections, particularly for variables such as GDP, are crucial for policymakers to make informed decisions about policies that can help sustain and improve the economy. GDP is a key indicator of a country's economic health and growth, and it is often used by policymakers to measure the overall performance of the economy and to determine the appropriate policy direction. By accurately projecting GDP growth, policymakers can plan for future spending, taxation, and investment decisions, and they can adjust policies to address economic challenges and opportunities. Projections for GDP growth can also help policymakers identify emerging trends and opportunities in the economy, such as new industries or changing consumer behavior, which can inform policy decisions.

Factors outside of the government's control, such as natural disasters and pandemics, can have a significant impact on a country's economic growth and can make it difficult to accurately predict future trends. The Covid-19 pandemic has had a particularly large impact on global economies, including Indonesia's. The pandemic has caused disruptions to supply chains, reduced consumer demand, and forced businesses to close or reduce their operations. As a result, Indonesia's GDP growth rate has decreased significantly in 2020 compared to the previous decade, which can make it challenging to accurately project future GDP growth rates based on historical data.

This paper describes an effort to develop nowcasting capacity in projecting Indonesia’s GDP. Using various kinds of data that are capable of transmitting economic conditions into projections, these data will be used in four different machine learning models in order to find out the model with the best individual performance, which is able to produce the closest projected value to the realization of GDP in the last quarter of 2022. The data used by the model will go through selection with certain criteria to determine economic indicators that have a strong relationship with GDP.

The paper follows pedagogical approach and structured as: Part 1 describes background, problem statement and objective of analysis. Part 2 provides literature on machine learning model that used in several countries/regions, including the machine learning concept used in the paper. Part 3 explains the

Should you require any further information, please contact us at:
ginanjar_utama@bi.go.id
nadira_firinda@bi.go.id

data and the methodology for using these models in the paper, followed by part 4 which presents our main results from different techniques used and discuss it. Part 6 concludes with summary of findings, limitations and recommendation for future research.

Literature Review

Overview of Machine Learning

Machine Learning is a subfield of artificial intelligence that develops method for teaching machines how to perform tasks by recognizing patterns and relationships in the provided data. Supervised and unsupervised learning are two main types of machine learning techniques. Supervised learning is used when the data has labeled examples, which means that each observation in the data is associated with a specific output or target variable that the model is trying to predict. The goal of supervised learning is to learn a function that maps the input variables to the output variable, based on the labeled data. Examples of supervised learning tasks include classification (predicting a discrete output variable) and regression (predicting a continuous output variable).

Unsupervised learning is a type of machine learning where the goal is to find patterns or structure in the data without the use of labeled examples. This means that the data is not labeled with an output variable or target, and the goal is to discover underlying patterns, relationships or clusters within the data. Examples of unsupervised learning tasks include clustering and dimensionality reduction (finding a lower-dimensional representation of the data).

Clustering is a type of unsupervised learning that involves grouping similar observations together based on their features or characteristics. There are many different clustering algorithms available, but they can generally be classified into two main categories: partition-based and hierarchical clustering. Partition-based clustering involves dividing the data into a fixed number of clusters, with each data point belonging to only one cluster. Hierarchical clustering, on the other hand, involves building a hierarchy of clusters, with each data point initially assigned to its own cluster, and then merging the clusters iteratively based on their similarity. There are also other types of clustering algorithms, such as density-based clustering and model-based clustering, that can be useful for different types of data and applications. By grouping similar observations together, we can gain insights into the data that may not be apparent from manual inspection or exploratory data analysis.

Economic Forecasting Using Machine Learning

Machine learning-based models have several benefits when it comes to macroeconomic projections. One of the main advantages of these models is their ability to incorporate all available data and can be tailored to focus on predicting specific events such as a recession. However, there are also several limitations associated with machine learning, such as the difficulty in interpreting the model's mechanisms or how it arrives at its predictions, the need for significant computational resources to handle large amounts of data, the dependence of variable selection on the chosen model, difficulties in solving large prediction problems, especially if many predictors are correlated, and the need for careful cross-validation to avoid overfitting.

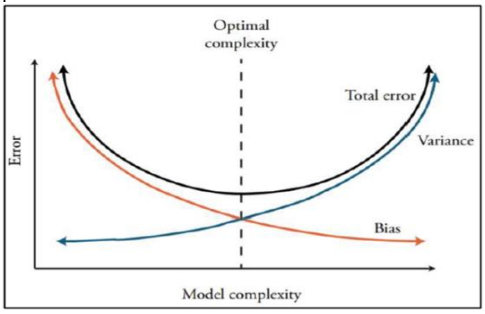
Should you require any further information, please contact us at:
ginanjar_utama@bi.go.id
nadira_firinda@bi.go.id

Variance-Bias Tradeoff

The use of the machine learning method as a forecasting tool is specifically designed to produce forecast results with the smallest error. The machine learning method decomposes forecasting error into two components, namely bias and variance. Bias is an error that results from inaccurate assumptions related to data, while variance is an error that results from model sensitivity due to changes in the underlying data.

In general, bias and variance will be affected by the complexity of the model which consists of 1) the number of variables in the model; 2) the number of parameters studied by the model; 3) the value of parameters specified at the start; and 4) the number of relationships between variables captured in the model. Variance and bias is a tradeoff where the more complex the model, the higher the variance and the lower the bias, and vice versa. As the ultimate goal, we will seek a model that has an optimal complexity that minimizes the total error that comes from variance or bias. Of course with the latest advancement of deep learning as a subset of machine learning, such as ChatGPT or other Large Language Models (LLM), the variance-bias tradeoff might not hold because of the emergent properties of complex systems.

Graph 1. Illustration of Variance-Bias Tradeoff



Source: Hall, A. S. (2018)

Nowcasting GDP using Machine Learning

This study refers to several studies that have previously been conducted on the economies of other countries or regions. Tiffin conducted a nowcasting of Lebanese GDP using the elastic net and random forest methods. Based on this research, Tiffin found that the elastic net method is better than the random forest method in nowcasting Lebanese GDP, and the combination of the two models provides the best accuracy.

Jung used machine learning models and additional data to improve one-quarter and one-year projections for seven middle- and high-income countries. The models used in this study are Elastic Net, SuperLearner, and Recurrent Neural Networks (RNN). The estimation results based on the machine learning model are able to provide better performance than the World Economic Outlook (WEO) projections.

Richardson uses New Zealand data for GDP nowcasting. The results show that the individual machine learning model is better than AR(1), and the combination model predicts even better. This study was continued by Richardson et al (2019) with 600 predictor data with various machine learning models, the results of the ML model are better than the benchmark AR(1), Dynamic Factor Model (DFM) and the official forecast.

Loermann also use machine learning artificial neural networks to nowcast the United States' GDP. They found that predicted results from Multi Layer Perceptron (MLP) was better than Dynamic Factor Model (DFM) benchmark model and Survey of Professional Forecaster (SPF).

Bolhuis, M., & Rayner, B. (2020) used Random Forest, Gradient Boosted Trees, Support Vector Machine, and the combined version of those three methods to nowcast Turkey's GDP. They found that the ensembled forecast method was better than the three methods individually, and the individual model itself was better than the benchmark Dynamic Factor Model.

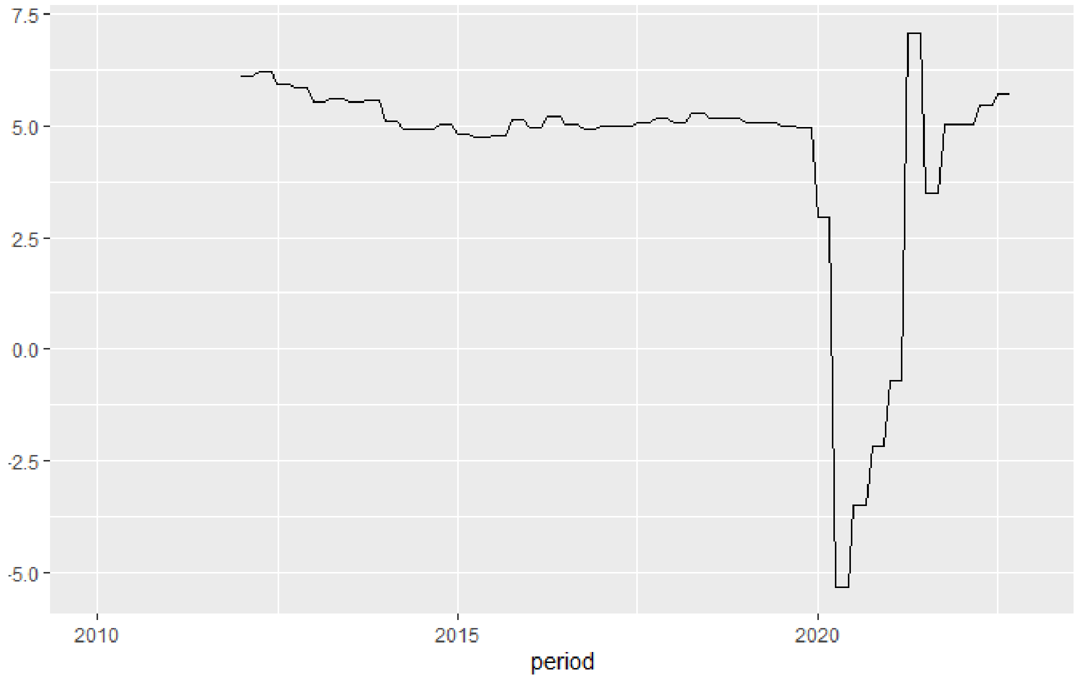
Hopp, D (2022) provides comparative results in nowcasting US GDP growth across three different volatile periods in US economic history. It examines the performance of 12 different methodologies including all the methods most commonly employed in nowcasting, as well as some of the most popular traditional machine learning approaches. The two best-performing methodologies in the analysis were found to be the long short-term memory neural networks (LSTM) and Bayesian vector auto regressian (BVAR).

Dauphin applies standar DFM and several machine learning algorithm to nowcast GDP growth across six European economies during normal and crisis times. The DFM tend to perform better during normal times while many of the ML methods perform strongly at identifying turning points. The performance of individual models greatly differs across sub-samples (time periods). The performances of individual DFM and ML models can also greatly differ across countries.They found that there is no one-size-fits-all method that would outperform the remaining methods in case of all countries under all circumstances.

Data and Methodology

As stated by Dauphin, The basic idea of nowcasting is to exploit a diverse set of timely information that is available before an official release of a target variable. As such, data selection and transformation is key to the success of nowcasting. In this research, the data used for the estimation is monthly indicator data both from a macroeconomic and financial perspective with an observation period from 2010 to 2022. There are a total of 34 variables, ranging from retail sales value to tax income, more of these variables are stated in Appendices. Later, the various data will be transformed into year-on-year percentage to prevent spurious relationship. It is important to note that the data used on a specific time range will have a one month lag. For example, nowcasting Q4 2022 GDP in November 2022 will use all data needed until October 2022 and so on.

Graph 2. GDP of Indonesia



Source: Central Bureau of Statistics (BPS)

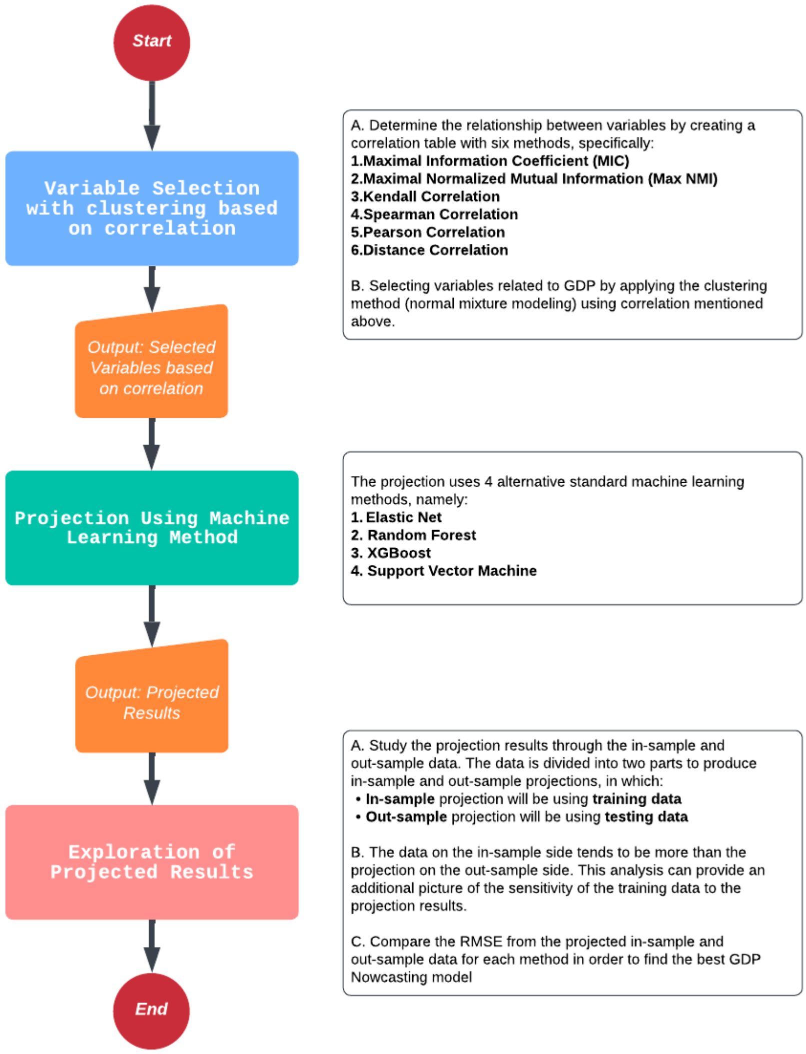
The historical data of GDP provides an essential input to the whole projection. According to the Central Bureau of Statistics, the realization of Indonesia's GDP was quite stable at the range of 5,0 to 7,5 (year-on-year percentage change) from 2010 until the year Covid-19 cases was first found. In 2020 until 2021, GDP dropped to the range of -5,0 to -0,6. Subsequently, Indonesia's GDP was back to its condition before the pandemic hits in the first until the third quarter of 2022.

The estimation will be divided into two parts, namely in-sample and out-of-sample projections. In-sample projections will use training data and out-of-sample projections will use testing data. Training data ranges between the first month of 2013 (2013M01) to December 2021 (2021M12), while testing data sets out from January (2022M01) to September 2022 (2022M09).

Future evaluations of the method chosen including using new data produced by Central Bureau of Statistics on the model periodically to asess the accuracy of the projection result. Evaluation has been done using the GDP realization data from the last quarter of 2022 (2022Q4).

The main ideas behind techniques applied in this study include: (i) determine the relationship between variables to agglomerate variables that have most linkages to GDP, (ii) project future values of GDP using the combined variables that have been selected based on their relationship, and (iii) asess the projection result through the in-sample and out-sample data. In depth, 34 high frequency indicators were exercised on 6 types of correlation (MIC, Max NMI, Kendall Correlation, Spearman Correlation, Pearson Correlation, and Distance Correlation). Afterward, based on their correlation to GDP they were selected by clustering using the clustering method called normal mixture modeling. The best combination of these variables will then be estimated iteratively using 4 machine learning models (Elastic Net, Random Forest, XGBoost, and Support Vector Machine) and complemented by ensemble model results to obtain predictions of GDP growth in the current period (quarter). Figure *1* describes the methodology applied using flow chart.

Figure 1. Methodology Flow Chart



Source: Internal

Furthermore, there will be an asessment conducted using the RMSE to measure the error given by projected in-sample and out-sample data for each method in order to find the best GDP Nowcasting Model. Consequently, future evaluations of the model is executed regularly to asess the accuracy of the projection result using the latest realization data.

Association/Relationship Measure

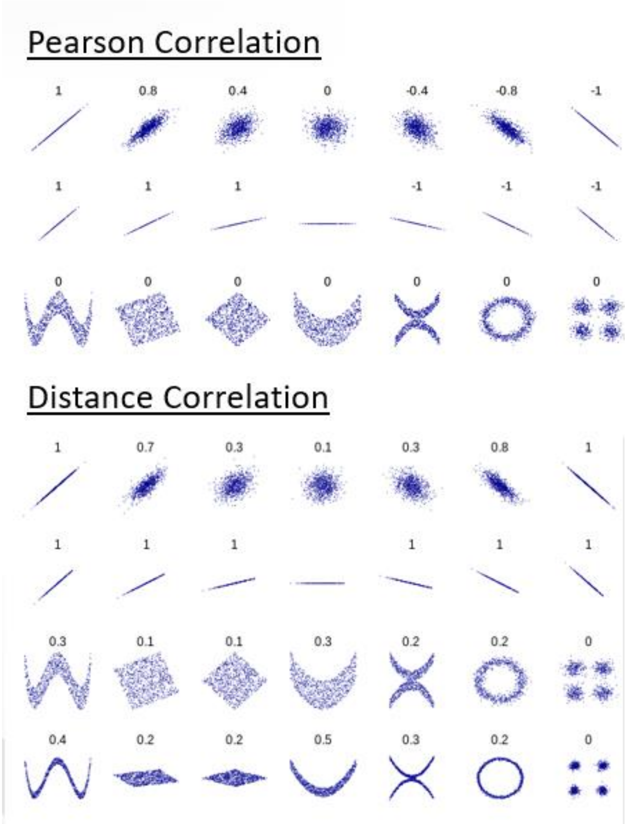
In determining the variables to be used in estimating using machine learning models, there are six association/relationship measures that will be used as variable selection criteria in this study. They are Maximal Information Coefficient (MIC), Maximal Normalized Mutual Information (Max NMI), Kendall Correlation. Spearman Correlation, Pearson Correlation, and Distance Correlation.

- Maximal Information Coefficient (MIC) is a measure of association that can capture functional and non-functional relationships between variables. For functional relationships, MIC can be associated with R^2 (coefficient of determination). It is based on the concept of mutual information, which measures the amount of information that two variables share. MIC finds the maximum value of the Pearson correlation coefficient between the variables over all possible partitions of the data. It ranges from 0 to 1, with higher values indicating a stronger relationship.

- **Mutual Information (MI)** is a measure of the relationship between two random variables that are sampled simultaneously. Specifically, MI measures how much information is communicated from one variable to another. Normalized Mutual Information (NMI) is a measure of MI that is normalized so that it is at a value between 0 and 1. Maximal **Normalized Mutual Information (Max NMI)** is normalized to take into account the entropy of the variables.
- **Kendall and Spearman Correlation** is a non-parametric test to measure the strength of dependency between two variables. There are both measures of association that are appropriate for ordinal data, which means that the values of the variables have a natural ordering. Kendall Correlation counts the number of concordant and discordant pairs of observations (i.e., pairs where the values of both variables are either both increasing or both decreasing, or one variable is increasing while the other is decreasing), while Spearman Correlation is based on the ranks of the observations rather than their raw values. Both measures range from -1 to 1, where values closer to 1 indicate a stronger association.
- **Pearson Correlation** is a parametric test to measure the strength of a linear relationship between variables, which means that it is sensitive to outliers and may not be appropriate for data that are not normally distributed. It is based on the covariance of the two variables divided by the product of their standard deviations. The assumptions needed are that the variables must be normally distributed, linear, and homoscedastic.
- **Distance Correlation** is a measure of the relationship between non-linear random variables. It goes beyond Pearson's correlation because it can spot more than linear associations and it can work multi-dimensionally. Distance correlation ranges from 0 to 1, where 0 implies independence between X & Y and 1 implies that the linear subspaces of X & Y are equal.

ginanjar_utama@bi.go.id
nadira_firinda@bi.go.id

Figure 2. Comparison of Pearson and Distance Correlation



Source: Wikipedia

The choice of which measure to use depends on the type of data we have, the nature of the relationship between the variables, and the particular problem we are trying to answer. We can sort and filter based on the strength of relationship and number of variables and continue to machine learning model to find the relationship patterns. We can also grouping the variables that has close relationship to GDP using the same characteristic by clustering method.

Clutering Method

Model-based clustering is a technique that aims to identify groups of data points within a dataset based on their underlying statistical distributions. This approach involves assuming that the data points in each group belong to a certain probability distribution, and then estimating the parameters of these distributions to identify the groups. One popular type of model-based clustering is the normal mixture model, which assumes that the data points in each group are normally distributed with a different mean and standard deviation. The normal mixture model assumes that the dataset is generated by a mixture of K normal distributions, each with its own mean and variance. The algorithm estimates the parameters of these K normal distributions, as well as the mixing proportions of the K groups, which determine the probability of each data point belonging to each group.

The estimation of the normal mixture model parameters is typically done using the Expectation-Maximization (EM) algorithm. In the E-step, the algorithm estimates the probability of each data point belonging to each group, based on the current estimate of the distribution parameters. In the M-step, the algorithm updates the estimates of the distribution parameters based on the probabilities estimated in the E-step. The E-M algorithm iteratively repeats these steps until convergence, resulting in the final estimates of the distribution parameters and the group assignments of each data point.

One advantage of model-based clustering, and specifically the normal mixture model, is that it can identify groups with different shapes and variances. Additionally, the algorithm can provide estimates of uncertainty in the group assignments and the distribution parameters, which can be useful for assessing the quality of the clustering results. However, the normal mixture model assumes that the data points in each group are normally distributed, which may not be appropriate for all types of data. Additionally, the performance of the algorithm can be sensitive to the initial parameter estimates and the number of groups specified.

Model Selection

The process of estimation and projection using machine learning involves a variety of model options. This study, however, narrows the selection to four commonly used models in similar research: Elastic Net, Random Forest, XGBoost, and Support Vector Machine. Each model will be generally described and can be further explored such as in James work. The tools and framework used in this research are R-based tools following Kuhn (2013,2020,2022) and Biecek, P (2018).

Elastic Net is a form of penalized regression by combining dimension reduction and variable selection in one step. This regularization method also provides robust results on potential correlations between predictors also known as the problem of multicollinearity. Multicollinearity occurs when two or more predictor variables in a linear regression model are highly correlated with each other, leading to unstable estimates of the regression coefficients and reduce the model's ability to generalize to new data. The trick is to add a penalty to the OLS regression so that all the coefficients are slightly downward biased which causes the addition of new information will not widen the predictions. This is known as regularization in machine learning terms.

$$\hat{\beta} = \arg\min_{\widehat{\beta}_J} \left\{ \sum_{i=1}^{n} \left(Y - X\hat{\beta} \right) + \mathit{Penalty}(\hat{\beta}) \right\}$$

Elastic net regression combines two penalties, namely LASSO (Least Absolute Shrinkage and Selection Operator) and Ridge. L1 regularization, also known as LASSO, adds a penalty term to the regression equation that penalizes the absolute value of the coefficients and encourages sparsity by setting some of the coefficients to zero. LASSO penalty apply the following equation:

$$\text{Penalty}(\hat{\beta}) = \lambda \sum_{j=1}^{p} |\beta|$$

Meanwhile, L2 regularization, also known as Ridge regression, adds a penalty term that penalizes the square of the coefficients and encourages the coefficients to be small but non-zero. Ridge penalty apply the following equation:

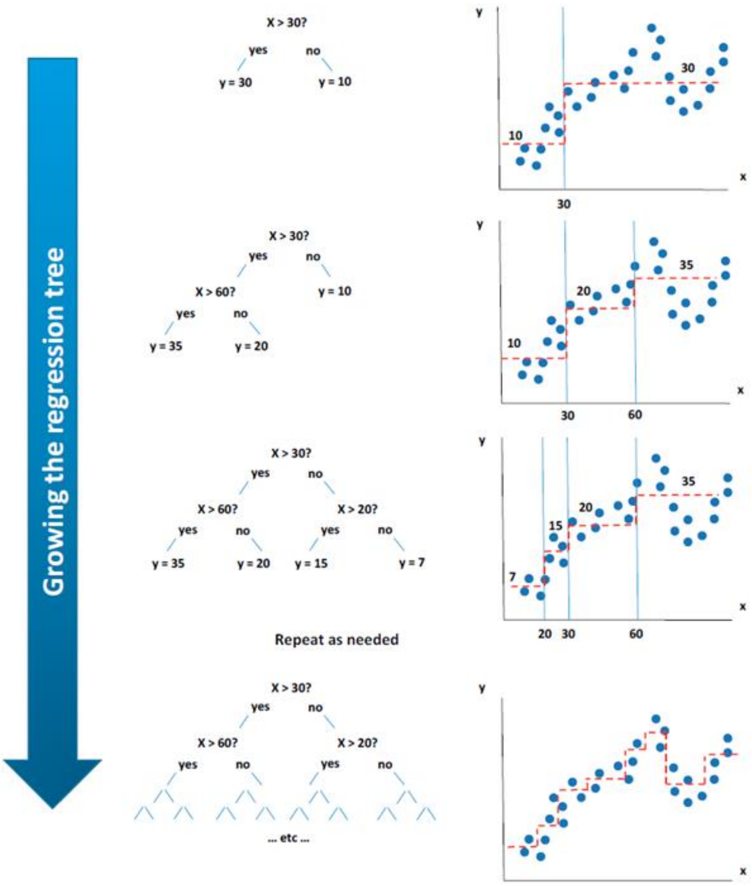
$$\text{Penalty}(\hat{\beta}) = \lambda \sum_{j=1}^{p} \left(\hat{\beta}\right)^2$$

By combining L1 and L2 regularization, Elastic Net can achieve a balance between sparsity and stability, which can improve the model's ability to generalize to new data.

Should you require any further information, please contact us at:
ginanjar_utama@bi.go.id
nadira_firinda@bi.go.id

A decision tree is a modeling technique that represents relationships in the form of a tree structure using iterative partitioning, where predictors are split into two sets starting with the initial division that minimizes the largest prediction error. When used for predicting continuous variables, a decision tree is referred to as a regression tree. Decision tree models offer several benefits, such as computational efficiency, the ability to model non-linear relationships, and robustness to missing data. However, the model is not well-suited for predicting linear relationships and is prone to overfitting. To address the overfitting issue in regression trees, Random Forest and Gradient Boosting methods are commonly used as alternative models.

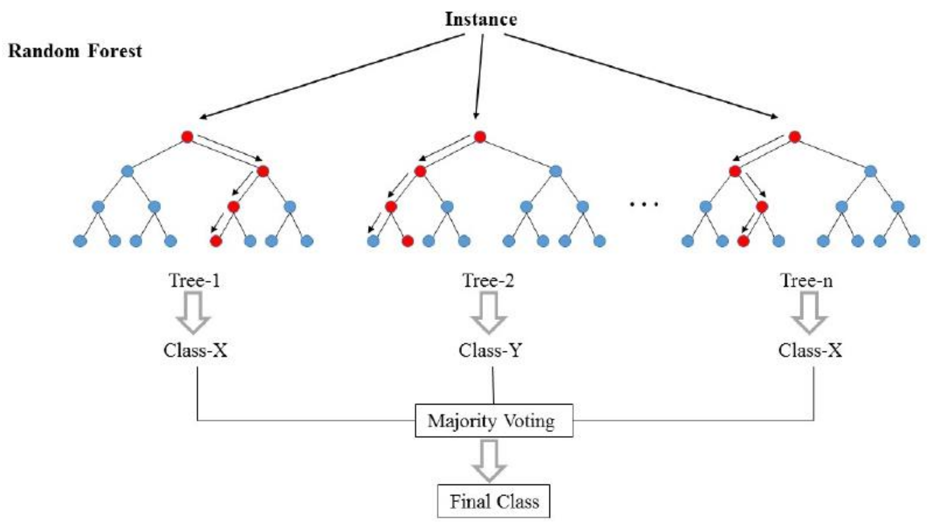
Figure 3. Illustration of Decision Tree



Source: Tiffin

Random Forest is an ensemble learning method that combines multiple decision trees to improve the accuracy and robustness of the model. Each decision tree is trained on a random subset of the training data and a random subset of the predictor variables, which helps to reduce overfitting and increase the diversity of the trees in the ensemble. The final prediction is then made by aggregating the predictions of all the trees in the forest. This bagging technique (bootstrap aggregation) helps to reduce overfitting and improve the accuracy and robustness of the model. Random Forest is a powerful and flexible algorithm that can be used for both classification and regression tasks.

Figure 4. Illustration of Random Forest

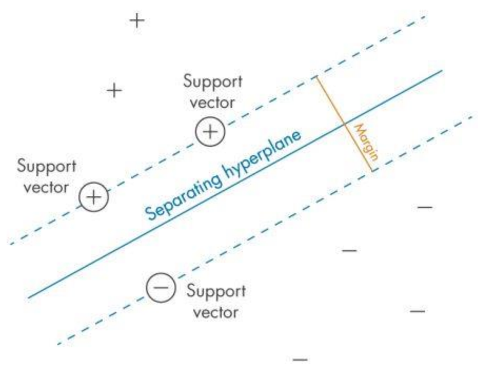


Source: Wikipedia

Meanwhile, XGBoost (eXtreme Gradient Boosting) is another ensemble learning method that is similar to Random Forest but uses a different approach to combine the individual models. XGBoost uses a gradient boosting framework, where each decision tree is trained to correct the errors of the previous trees, in order to improve the overall accuracy of the ensemble. XGBoost is known for its speed, scalability, and ability to handle large datasets with high-dimensional features.

Support Vector Machine (SVM) is a non-parametric technique that aims to identify hyperplanes that maximize the distance between classes in the feature space while ensuring that the perpendicular distance between two adjacent points of different classes is maximized. SVM is advantageous because it is effective in high dimensions, especially when the number of dimensions exceeds the number of samples, it efficiently utilizes memory, and it can employ various kernel functions for the decision function. However, the model may be prone to overfitting, which can be addressed by carefully selecting kernel functions and utilizing cross-validation. This is particularly important when the number of features is greater than the number of samples.

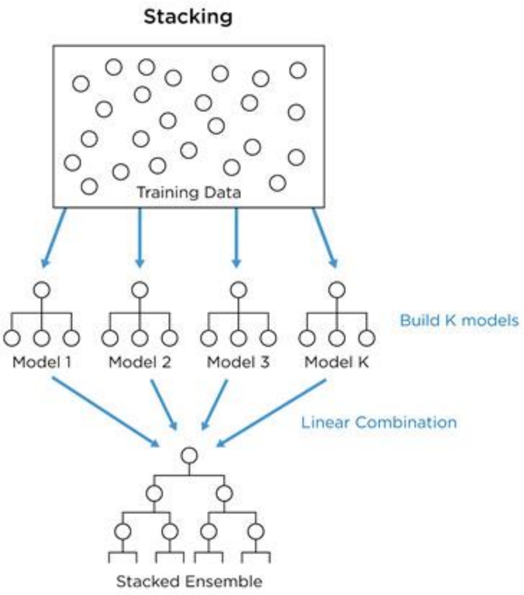
Figure 5. Illustration of Support Vector Machine



Source: de.mathworks.com/discovery/support-vector-machine.html

Instead of relying on individual models that produce inconsistent or inconclusive predictions, this study also proposes the possibility to use of ensemble models that combine predictions from multiple model results. The ensemble model follows the Diversity Prediction Theorem introduced by Page where the collective/crowd error of various models will be equal to the average individual error minus the prediction diversity. Therefore, when the diversity in a group is large, the collective error is small.

Figure 6. Illustration of Ensemble Model



Source: https://www.kdnuggets.com/2019/09/ensemble-learning.html

Results and Discussion

Evaluation of Model Accuracy

To determine the best machine learning model for nowcasting Indonesia's GDP, a comparison was made of the RMSE of each model and the performance of each model presented in a line graph-like visualization.

Graph 3. RMSE of in-sample and outsample projections

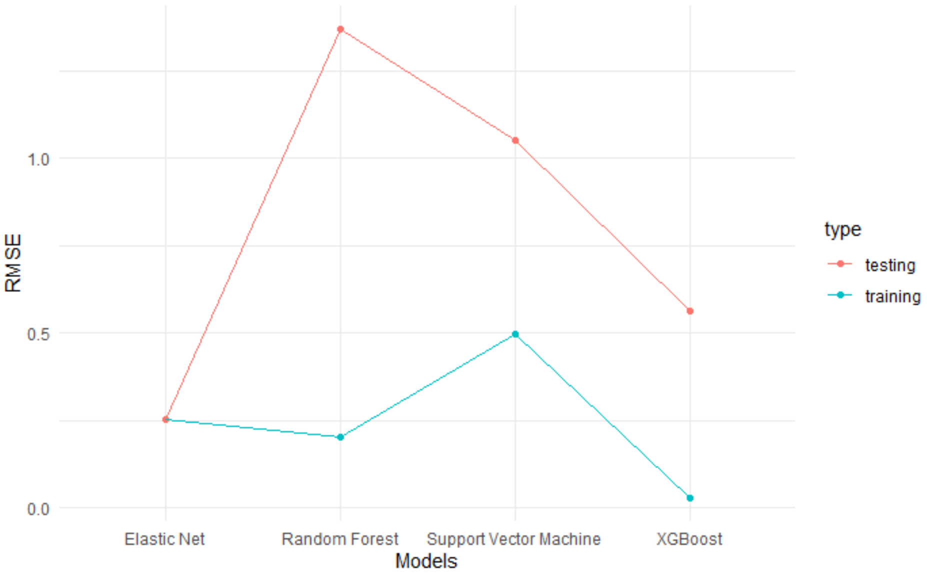


Table 1. RMSE Value for Training and Testing Data

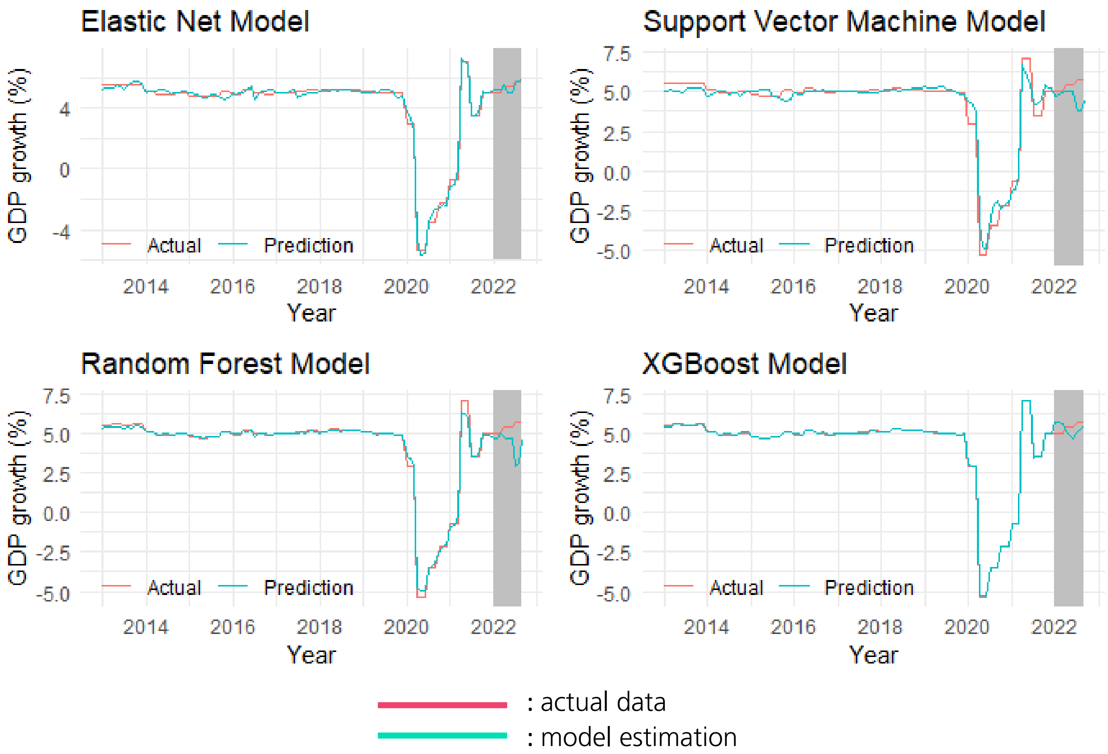
Model	Training	Testing
Elastic Net	0.2541254	0.2531369
Random Forest	0.2041514	1.3689247
XGBoost	0.0296590	0.5628105
Support Vector Machine	0.4971251	1.0499439

Source: Internal Calculation

The line graph shows that the RMSE values from the experiments using the training and testing data by the Elastic Net model are quite close. This phenomenon is followed by the RMSE values for both variance of data issued by XGBoost, followed by SVM and Random Forest respectively. In line with the line graph, Elastic Net is proven to be the model with the smallest difference in RMSE values between training and testing data (0.001), this certainly shows that the projected value generated using the training data is quite close to the projected value of the testing data. In the Random Forest, the difference in the RMSE value between the experiment and the training and testing data is 1.165. Followed by XGBoost with a difference in RMSE value of 0.053 and SVM of 0.522.

As stated before in Data, the exploration uses training data from January 2013 (2013M01) to December 2021 (2021M12) and testing data from the first month of 2022 (2022M01) to September 2022 (2022M09). We have selected ten variables as input in order to find out which type of model is capable of providing the best projection of GDP.

Graph 4. In-sample and Out-sample Projection from Each Models

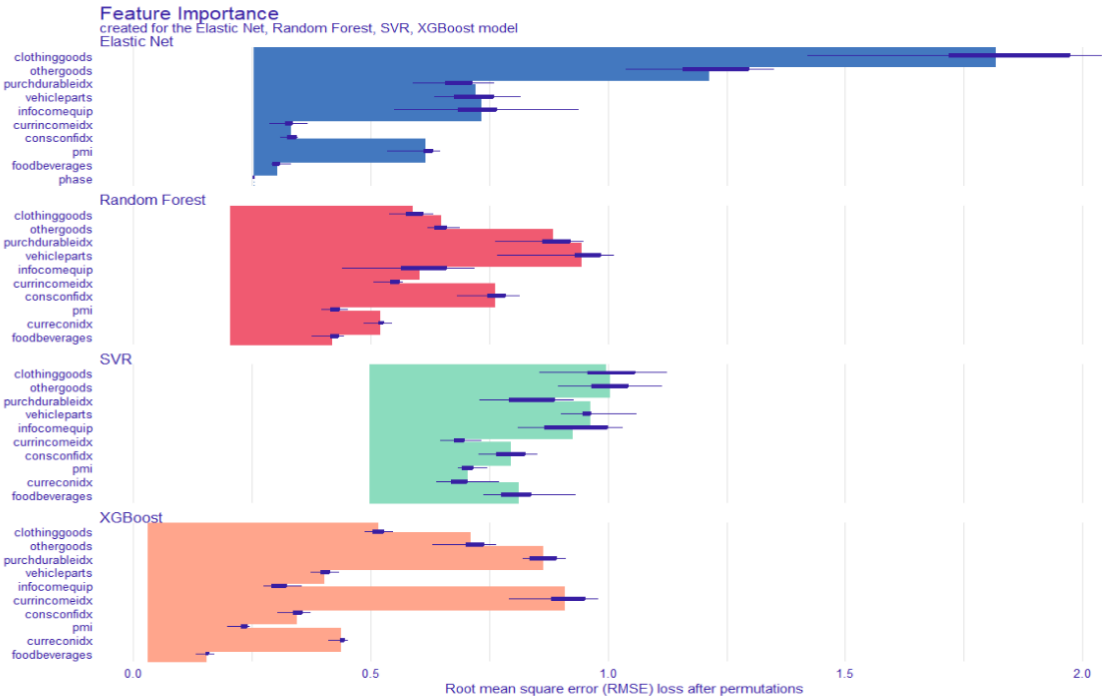


Source: Internal Calculation

Feature Importance

Feature importance aka Permutation-based variable importance is a method for determining the relative importance of predictor variables in a machine learning model. The idea behind this method is to measure how much the performance of the model degrades when the values of a particular predictor variable are randomly permuted, which can help identify which variables are the most important for accurate predictions. If a predictor variable is important for accurate predictions, then randomly permuting its values should lead to a noticeable decrease in model performance. On the other hand, if a variable is not important, then permuting its values should have little effect on model performance. One advantage of permutation-based variable importance is that it is model-agnostic, meaning that it can be applied to any type of machine learning model.

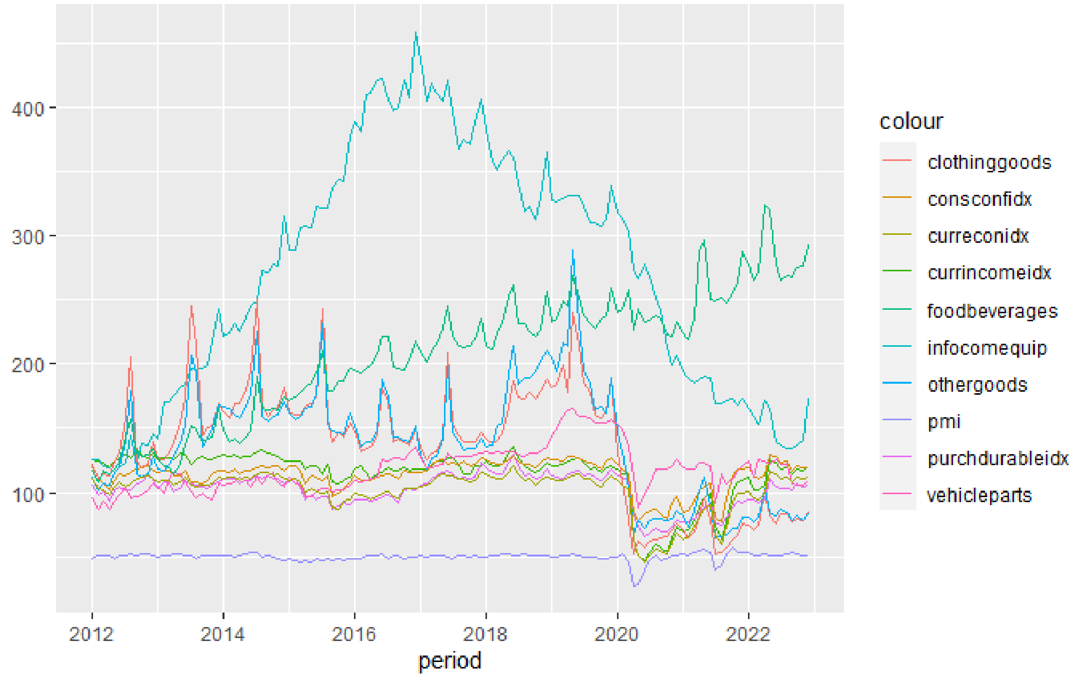
Graph 5. Feature Importance from The Models



Source: Internal Calculation

Based on the Feature Importance graph, the XGBoost model produces the smallest RMSE loss after permutation. Based on the feature importance of each variable, the GDP nowcasting results are most influenced by the consumption aspect, namely Clothing and Other Goods. Some of the variables assessed also have a significant role in GDP nowcasting including the Purchase of Durable Goods Index (purchdurableidx), Information and Communication equipment (infocomequip) and Index of Current Income (currincomeidx).

Graph 6. Selected input variables in the final model (level)



Source: Internal Calculation

Graph *6* shows the year on year growth on the selected variables based on six different criterias from 2012 to 2022. It is implied that Clothing and Other Goods have the same pattern over the years, similar to Index of Current Income, Consumer Confidence Index, and Purchase of Durable Goods Index. Furthermore, Index of Current Income seems to have constant, patterned increase from 2012 with the execption from 2021. While Infocomequip appears to have the highest increase of all variables from around 2017 before dropping gradually in 2018 henceforth.

Nowcasting Result and Model Explanation

Table 2. Nowcasting GDP growth for Q4 2022 in Indonesia

Indicator Data Periode	Elastic Net	Random Forest	XGBoost	Support Vector Machine (SVM)	Average Ensembled
October	5.384883	4.766341	4.868336	4.974105	4.998416
November	5.292687	4.854379	4.689212	4.965604	4.950471
December	5.188386	4.898498	4.886504	4.948184	4.980393
Average	5.288652	4.839739	4.814684	4.962631	4.976427

Source: Internal Calculation

So far, the key findings from the models and nowcasting results include the following:

- The combination between Consumer Confidence Index (consconfidx), Index of Current Income (currincomeidx), Purchase of Durable Goods Index (purchdurableidx), Current Economy Index (curreconidx), PMI (pmi), and real retail sales indicators such as: Information and

Communication equipment (infocomequip), Food and Beverages (foodbeverages), Vehicle Parts (vehicleparts), Clothing (clothinggoods) and Other Goods (othergoods) have constantly high contributions value in regards to GDP of Indonesia.

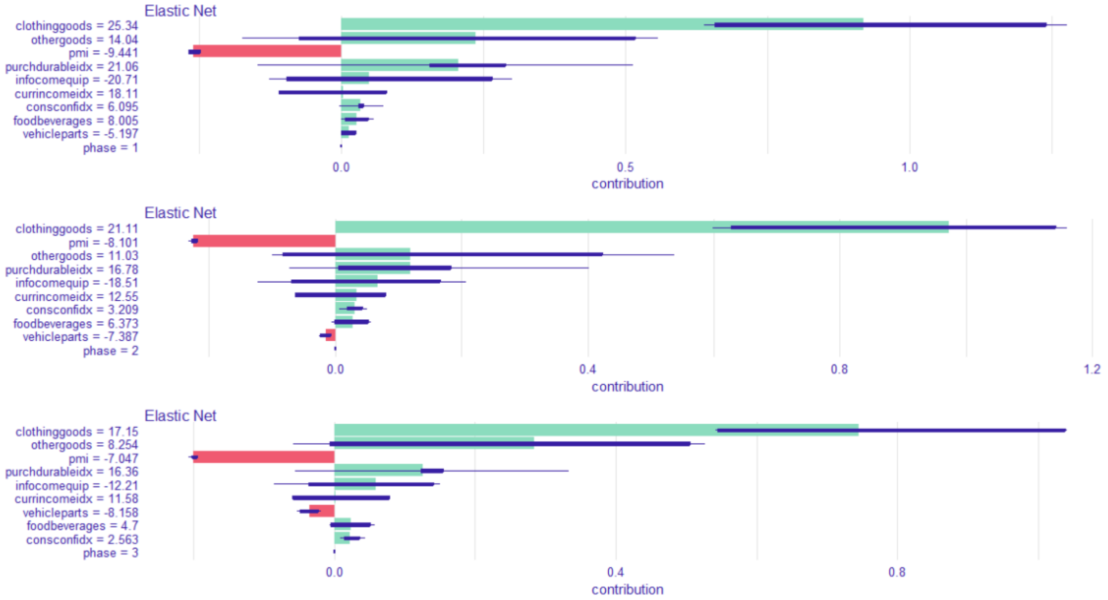
- Based on the performance of individual models, Elastic Net and XGBoost have RMSE values of residuals which are quite low compared to Random Forest and SVM in projecting Indonesia's GDP. it can be concluded that Elastic Net and XGBoost provide the best performance compared to Random Forest and SVM as an individual model capable of projecting Indonesia's GDP.
- The average projection of GDP in Indonesia over the course of the last quarter of 2022 from Elastic Net, and Random Forest models are 5.29 and 4.84 consecutively. While the GDP projected by XGBoost and SVM is 4.81 and 4.96 respectively. Furthermore, the average projection that ensembles multiple results from Elastic Net, Random Forest, XGBoost and SVM is 4.98.

It is important to not only evaluate the accuracy of the model's projections, but also to understand and explain how the model arrived at its predictions. This transparency and accountability can be achieved through interpretation methods such as Shapley Value and Breakdown Profile decomposition, as recommended by Molnar and Biecek, P. & Burzykowski, T. (2021).

- **Shapley Value** is used to perform an averaging of the value of a variable's attribution over all (or a large number of) possible orderings. This is derived from game theory which we can see the contribution of players in a cooperative game (coalition). Shapley values provide a uniform approach to decompose a model's predictions into contributions that can be attributed additively to different explanatory variables. An important drawback of Shapley values is that they provide additive contributions (attributions) of explanatory variables. If the model is not additive, then the Shapley values may be misleading. Along with the GDP projection, the tendency of the contribution of the variables to GDP can also be analyzed. In this case, using the data between October to December 2022, some particular variables have constant importance in order.

Shapley Value from Elastic Net

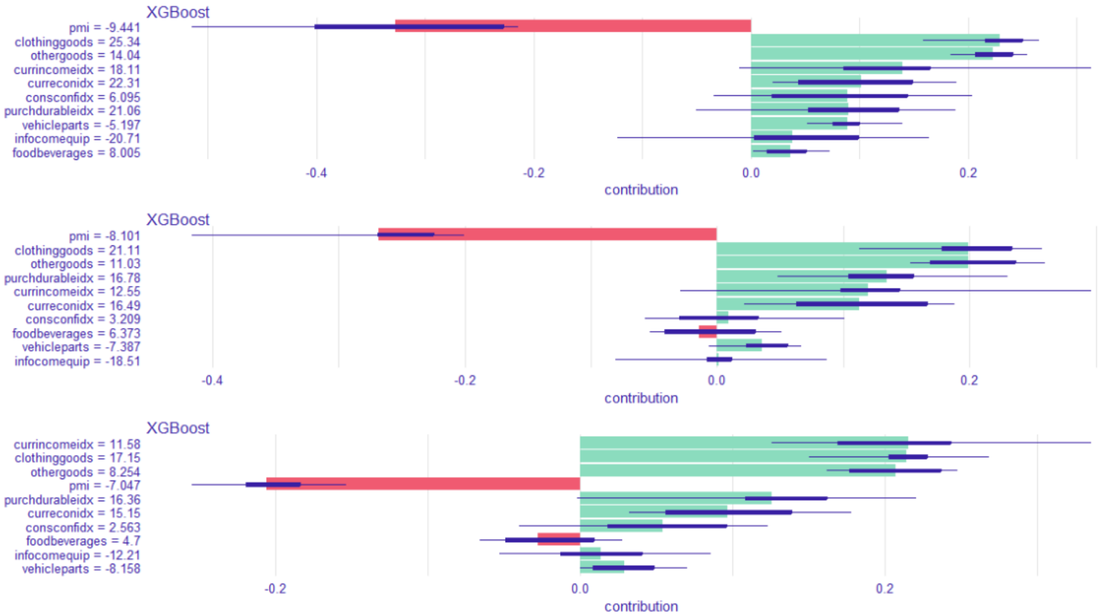
Graph 7. Shapley Value with The Period of October to December 2022



Source: Internal Calculation

Shapley Value from XGBoost

Graph 8. Shapley Value with The Period of October to December 2022



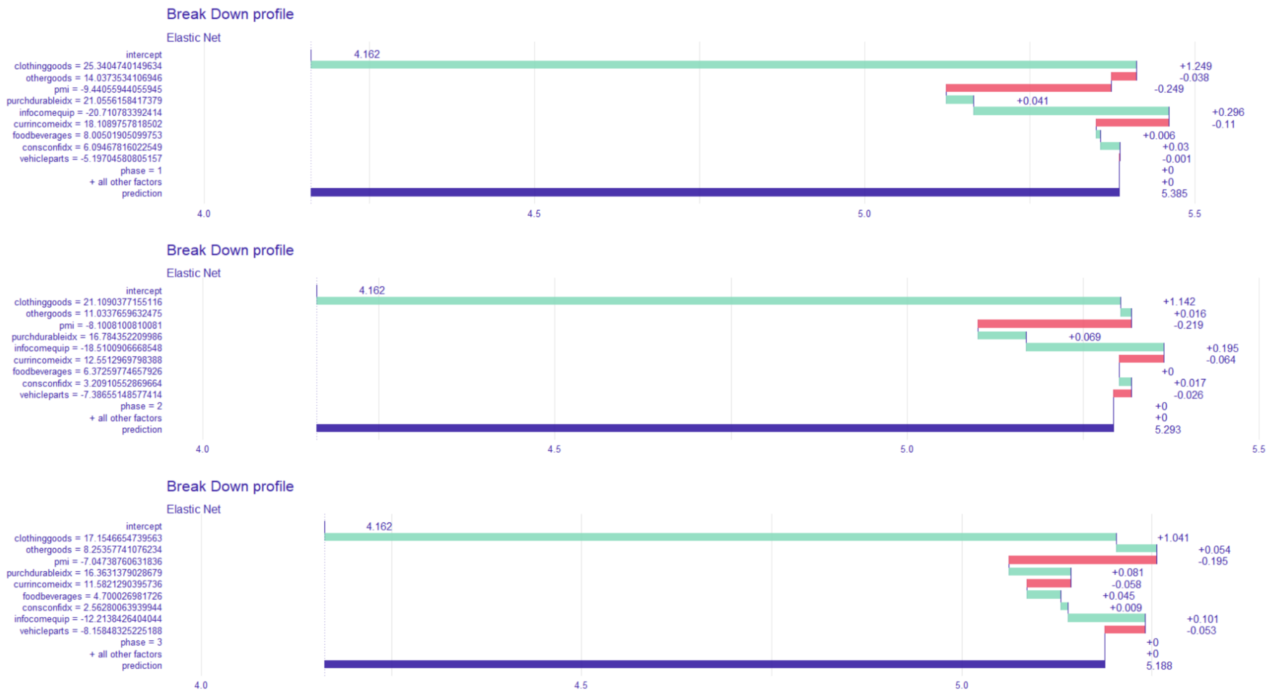
Source: Internal Calculation

Elastic Net and XGBoost, as the best models elected to project GDP so far, have simliar results on the variables' contribution in shaping Q4 2022 GDP. In general, GDP consists of 10 essential variables, that are Index of Current Income (currincomeidx), Purchase of Durable Goods Index (purchdurableidx), Information and Communication equipment (infocomequip), Consumer Confidence Index (consconfidx), Clothing Goods (clothinggoods), Other Goods Index (othergoods), Parts of Vehicles (vehicleparts), Food and Beverages (foodbeverages), PMI, and Index of Current Economy (curreconidx) where both Clothing and Other Goods Index are considered to have the biggest impact on Q4 2022 GDP compared to other variables in projections generated by the said models respectively. Meanwhile, PMI is a variable that is inspected to have negative contribution to Q4 2022 GDP. Following Clothing and Other Goods Index are Purchase of Durable Goods Index, Information and Communication equipment and Index of Current Income, at which most cases include Purchase of Durable Goods Index has bigger importance than the other two variables.

- Using **Breakdown Profile**, we also evaluate which variables contribute most to the result. The intercept represents the mean value that can be interpreted as an estimate of the expected value of the model's predictions over the distribution of all explanatory variables. The green and red bars indicate, respectively, positive and negative changes in the mean predictions (contributions attributed to explanatory variables). In general, using the data from October to December 2022 both Elastic Net and XGBoost has the same result with Clothing and Other Goods have the most contribution for nowcasting GDP in Q4 2022.

Break Down from Elastic Net

Graph 9. Breakdown Profile with The Period of October to December 2022

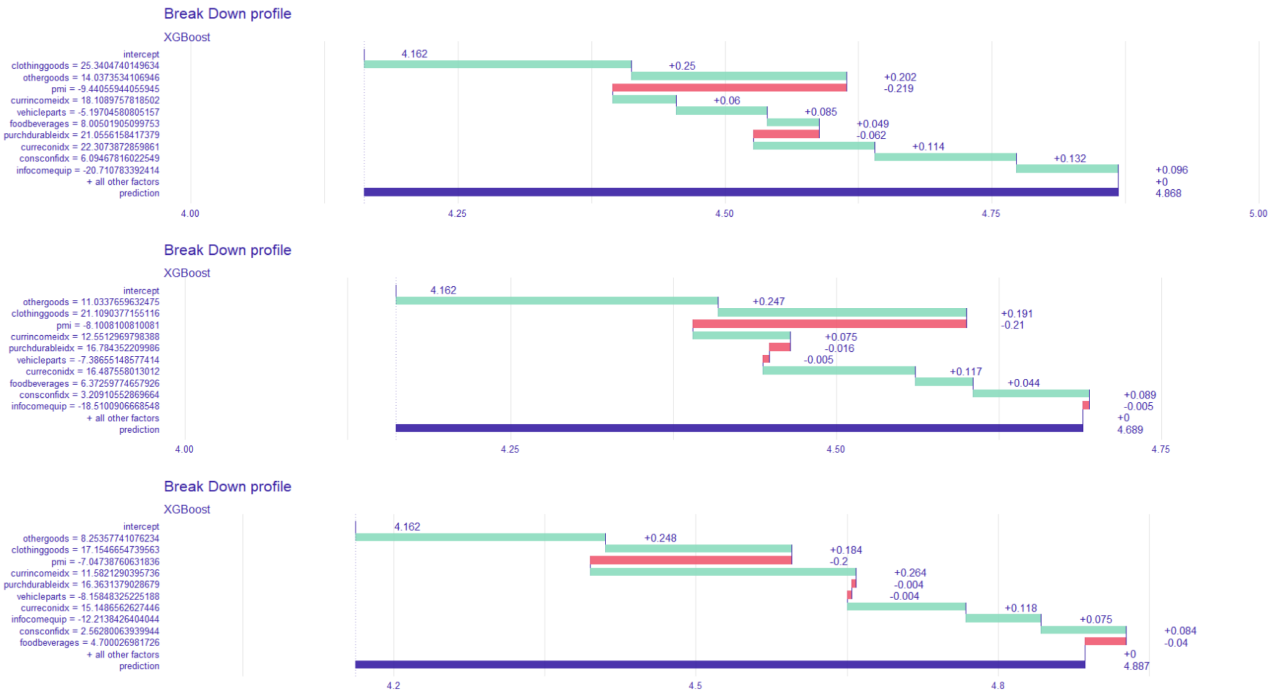


Source: Internal Calculation

From the Breakdown Profile graph, it is shown that between October to December 2022, Clothing Goods increase the value of prediction from the mean by around 1,041 to 1,249, making it as the variable with the highest contribution to GDP over the others. Following Clothing Goods is Purchase of Durable Goods Index that has the range of increasing the predition by 0,101 to 0,269. Meanwhile, in line with the results of assessment brought by Shapley Value, PMI is shown to have a negative contribution in regards to other variables with decreasing the projection at intervals -0,249 to -0,195.

Break Down from XGBoost

Graph 10. Breakdown Profile with The Period of October to December 2022



Source: Internal Calculation

With lesser impact than the result shown by Elastic Net, the breakdown profile from XGBoost model implies that between October to December 2022, Clothing Goods increase the probability of prediction by around 0,248 to 0,250, still the variable with the highest contribution to GDP. On contrary with Elastic Net, Other Goods is the next variable with upmost contribution to GDP with the intervals of 0,184 to 0,202. Meanwhile, in line with the results of assessment brought by Shapley Value, PMI is shown to have a bigger negative contribution in the observation compared to other variables with plummeting the projection between -0,2 to -0,219.

Conclusion

According to our findings, the projected GDP growth for Q4 2022 in Indonesia ranges from 4.81% to 5.29%, with an average of 4.98% using an ensemble method. The Root Mean Squared Error (RMSE) ranges from 0.03 to 0.49 for the training data, which consists of monthly data from January 2010 to December 2021. The RMSE ranges from 0.25 to 1.36 for the testing data, which includes monthly data from January to September 2022. In addition, to make the projection more transparent and interpretable, we performed a Shapley value decomposition and Breakdown Profile decomposition. Our analysis also indicates that some indicators, such as the Current Incomes Index, Consumer Confidence Index, Purchase of Durable Goods Index, PMI, and real retail sales indicators, including Information and Communication equipment, Food and Beverages, Vehicle Parts, Clothing, and Other Goods, have consistently contributed significantly to the GDP of Indonesia. Our results demonstrate that our model can predict GDP growth accurately in both normal and pandemic periods.

Our machine learning model for nowcasting GDP growth has limitations due to our limited knowledge and resources. However, as the development of machine learning and data digitization continues, more opportunities to improve model performance will arise. Continual evaluation and development of models will lead to more accurate and accountable macroeconomic projections that will aid policy formulation and transmission. To ensure responsible and transparent AI adoption, organizations should consider the four areas of Responsible AI principles: internal governance structures and measures, determining the level of human involvement in AI-augmented decision-making, operations management, and stakeholder interaction and communication. Clear roles and responsibilities should be established, and a risk-based approach should be used to minimize bias in data and models used for AI. Users should be included in the process and provided with clear communication about AI policies.

Reference

Biecek, P. & Burzykowski, T. (2021). *Explanatory Model Analysis: Explore, Explain, and Examine Predictive Models*. Chapman and Hall/CRC, New York

Biecek, P (2018), *DALEX*: *Explainers for Complex Predictive Models in R. Journal of Machine Learning Research* 19(84):1–5, 2018.

Bolhuis, M., & Rayner, B. (2020). *Deus Ex Machina? A Framework for Macro Forecasting with Machine Learning*. IMF Working Papers, WP/20/45. International Monetary Fund.

Dauphin, Jean-Francois and Dybczak, Kamil and Maneely, Morgan and Taheri Sanjani, Marzie and Suphaphiphat, Nujin and Wang, Yifei and Zhang, Hanqi (2022). *Nowcasting GDP - A Scalable Approach Using DFM, Machine Learning and Novel Data, Applied to European Economies*. IMF Working Paper No. 2022/052

Hall, A. S. (2018). *Machine Learning Approaches to Macroeconomic Forecasting*. The Federal Reserve Bank of Kansas City Economic Review, 103(63), 2.

Hopp, D (2022). *Benchmarking econometric and machine learning methodologies in nowcasting*. UNCTAD Research Paper No. 83

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning*. New York: Springer.

Jung, J. K., Patnam, M., & Ter-Martirosyan, A. (2018). *An Algorithmic Crystal Ball: Forecasts Based on Machine Learning*. IMF Working Paper 18/230.

Kuhn, M, and K Johnson. 2013. *Applied Predictive Modeling*. Springer.

Kuhn, M, and K Johnson. 2020. *Feature Engineering and Selection: A Practical Approach for Predictive Models*. CRC Press.

Kuhn, M, and Silge, J. 2022.*Tidy Modeling with R: A Framework for Modeling in the Tidyverse*, O'Reilly Media, Inc.

Loermann, J. & Maas, B. (2019). *Nowcasting US GDP with Artificial Neural Networks*. MPRA Paper 95459, University Library of Munich, Germany.

Molnar, C. (2021). *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. https://christophm.github.io/interpretable-ml-book/

Richardson, A., & Mulder, T. (2018). *Nowcasting New Zealand GDP Using Machine Learning Algorithms*," Mimeo.

Richardson, A., Mulder, T., Vehbi, T. (2019*). Nowcasting GDP using Machine Learning Algorithms: A Real-Time Assessment*. Discussion Paper, 2019-03. Reserve Bank of New Zeland.

Page, S.E. (2010). *Diversity and Complexity (Primers in Complex Systems)*. Princeton University Press

Tiffin, A. (2016). *Seeing in the Dark: A Machine-Learning Approach to Nowcasting in Lebanon*. IMF Working Paper 16/56.

Appendices

Nowcasting Variables of Indonesia

Code	Variable	Frequency	Data Lag	Unit
gdprl	GDP of Indonesia	Monthly	1 Month	YoY Percentage (%) Change
retailsales	Retail Sales	Monthly	1 Month	YoY Percentage (%) Change
vehicleparts	Vehicle Parts	Monthly	1 Month	YoY Percentage (%) Change
foodbeverages	Food and Beverages	Monthly	1 Month	YoY Percentage (%) Change
autofuels	Fuels of Automotive	Monthly	1 Month	YoY Percentage (%) Change
infocomequip	Information and Communication equipment	Monthly	1 Month	YoY Percentage (%) Change
hholdequip	Household Equipment Index	Monthly	1 Month	YoY Percentage (%) Change
recreationgood	Recreation Goods	Monthly	1 Month	YoY Percentage (%) Change
othergoods	Other Goods	Monthly	1 Month	YoY Percentage (%) Change
clothinggoods	Clothing Goods	Monthly	1 Month	YoY Percentage (%) Change
mobilsales	Automobile Sales	Monthly	1 Month	YoY Percentage (%) Change
motorsales	Motorcycle Sales	Monthly	1 Month	YoY Percentage (%) Change
prod_motor	Motorcycle Production	Monthly	1 Month	YoY Percentage (%) Change
pmi	Purchasing Managers Index	Monthly	1 Month	YoY Percentage (%) Change
farmertradeidx	Index of Farmer Trade	Monthly	1 Month	YoY Percentage (%) Change
idx_ihsg	Composite Stock Price Index	Monthly	1 Month	YoY Percentage (%) Change
consconfidx	Consumer Confidence Index	Monthly	1 Month	YoY Percentage (%) Change
curreconidx	Current Economy Index	Monthly	1 Month	YoY Percentage (%) Change
consexpctidx	Consumer Expectation Index	Monthly	1 Month	YoY Percentage (%) Change
currincomeidx	Current Income Index	Monthly	1 Month	YoY Percentage (%) Change
jobavailidx	Job Availibility Index	Monthly	1 Month	YoY Percentage (%) Change
purchdurableidx	Purchase of Durable Goods Index	Monthly	1 Month	YoY Percentage (%) Change
rtgstx	Real Time Gross Settlement (RTGS) Transaction Value	Monthly	1 Month	YoY Percentage (%) Change
skntx	National Clearing System Transaction Value	Monthly	1 Month	YoY Percentage (%) Change
marketcap	Market Capital	Monthly	1 Month	YoY Percentage (%) Change

Should you require any further information, please contact us at:
ginanjar_utama@bi.go.id
nadira_firinda@bi.go.id

idx_lq45	LQ45 Stock Index	Monthly	1 Month	YoY Percentage (%) Change
idx_basic_ind	Basic Industry Stock Index	Monthly	1 Month	YoY Percentage (%) Change
idx_infr	Infrastructure Index	Monthly	1 Month	YoY Percentage (%) Change
idx_finance	Financial Index	Monthly	1 Month	YoY Percentage (%) Change
reserve	Reserve	Monthly	1 Month	YoY Percentage (%) Change
exrpl	Exchange Rate of Indonesia	Monthly	1 Month	YoY Percentage (%) Change
crude_oil	Crude Oil	Monthly	1 Month	YoY Percentage (%) Change
nontaxincome	Income from Non-Tax	Monthly	1 Month	YoY Percentage (%) Change
taxincome	Income from Tax	Monthly	1 Month	YoY Percentage (%) Change
l1prod_motor	Lag 1 month Motorcycle Production Volume	Monthly	1 Month	YoY Percentage (%) Change