



Predicting Indonesia's Gross Domestic Product (GDP): A Comparative Analysis of Regression and Machine Learning Models

Christopher Kevin Widjaja^{1*}, Soo Huei Ching², Bernadette Lawardi³

^{1,2} *School of Mathematical and Computer Sciences, Heriot-Watt University Malaysia, 62200 Malaysia*

³ *PT NielsenIQ Services Indonesia, 12920 Jakarta, Indonesia*

*Corresponding author: ck2004@hw.ac.uk

Received 29 July 2023

Accepted 5 September 2023

Published 25 October 2023

Abstract

This research paper presents an analysis of Indonesia's quarterly Gross Domestic Product (GDP) growth spanning a significant 13-year period, from the first quarter of 2010 to the fourth quarter of 2022. The study focuses on utilizing four key economic indicators to gain insights into the country's economic performance during this timeframe. To develop accurate predictive models, we utilize Multiple Linear Regression (MLR), K-Nearest Neighbours (K-NN), and Artificial Neural Network (ANN) approaches. The models are compared based on performance metrics, including the Mean Absolute Error (MAE) and the Root Mean Squared Error (RMSE). Our findings indicate that the MLR model outperforms the machine learning models in forecasting Indonesia's GDP. The MLR model is approximately 93.3% better than the K-NN model and approximately 64.7% better than the ANN model based on the RMSE values. This suggests that a simpler and more explainable model, such as MLR, suffices to provide meaningful and interpretable results. The paper's insights are valuable to economists, policymakers, and researchers, offering a practical and understandable means to predict Indonesia's economic trajectory.

Keywords: Artificial Neural Network (ANN), Gross Domestic Product (GDP), K-Nearest Neighbours (K-NN), Multiple Linear Regression (MLR), Predictive Models.

RESEARCH ARTICLE

1. Introduction

The Republic of Indonesia, a Southeast Asian archipelago strategically located between Asia and Oceania, is the fourth most populous country with over 277 million people spread across 38 provinces and more than 17,000 islands. Its growing economy, driven by a large consumer market, urbanization, and government policies, is projected to become the world's fourth-largest economy by 2045, supported by investments in infrastructure and natural resources.

According to the Central Agency on Statistics, known as Badan Pusat Statistik (BPS), a non-departmental government institute responsible for conducting statistical surveys in Indonesia, Gross Domestic Product (GDP) and Gross Regional Domestic Product (GRDP) are pivotal indicators of an economy's overall health and performance. GDP measures the total value of goods and services produced within a country in a specific time period, typically a year, and is widely used as a benchmark

for comparing the economic performance of different countries (Callen, 2012). In contrast, GRDP assesses the total value of goods and services produced at a regional or sub-national level instead of the national level, offering detailed insights into specific areas (BPS, 2020).

In this study, our objective is to assess Indonesia's growing economy by leveraging the Gross Domestic Product (GDP) data based on a constant market process. We use a dataset accessible online through the BPS website, which comprises quarterly data on Indonesia's total GDP and the GRDP of its 38 provinces from 2010 to 2022. To construct predictive models, we focus on four key economic indicators contributing to the Indonesian GDP: Net Export of Oil and Non-Oil Commodities (NE), Household Consumption (HC), Government Spending (GS), and Inventory Changes (IC). Through our analysis, we aim to gain insights into the economic trends and forecast future developments, contributing to a deeper understanding of Indonesia's economic landscape.

The layout of the paper is outlined as follows. In Section 2, we provide a brief literature review on GDP modelling. Section 3 discusses the methodology and the performance metrics used. In Section 4, we present the data descriptive statistics and highlight the significant findings. Finally, in Section 5, we offer conclusions and further discussions based on the study's outcomes.

2. Literature Review

Numerous studies have analyzed various aspects of GDP using various models. The utilization of modern tools and techniques enables researchers to deal with significant amounts of data accurately, leading to better decision-making, predictive analysis, and pattern recognition. These insights provide a deeper understanding of the factors that affect GDP in certain countries or regions and highlight areas that encourage economic growth.

Nwabueze (2009) investigated Nigeria's GDP and personal consumption expenditure relationship using regression analysis. The findings showed no significant effect on personal consumption expenditure as Nigeria's GDP increased, indicating a limited explanatory power of GDP on personal consumption expenditure. This may be due to GDP only explained about 3.5% of Nigeria's personal consumption expenditure.

Desai and Bhatia (2016) highlighted the use of a multivariable regression model to forecast India's GDP growth, linking it to key drivers such as CPI inflation, manufacturing and services purchasing managers' indexes, interest rates, and crude oil prices. The relationships between the key drivers and the GDP were observed using linear regression models.

Dudu and Moscu (2016) employed multiple linear regression to evaluate Romania's GDP, identifying factors like household consumption, public consumption, and gross investments that influenced it. The study suggested Romania's overreliance on imports, emphasizing the need to bolster domestic demand.

Stanić and Račić (2019) discussed the application of multiple linear regression analysis in macroeconomic research, focusing on evaluating the effects of various macroeconomic factors on GDP in Bosnia and Herzegovina from 2005 to 2018. The study utilized six economic indicators, namely foreign direct investments, imports, exports, growth rate, unemployment, and inflation rate. The findings revealed that among these factors, Bosnia and Herzegovina's imports had the most significant effect on GDP, followed by Direct Foreign Investment (FDI) and exports. As a result, the study's outcomes provided valuable insights for policymakers, enabling them to make informed decisions to promote economic growth, particularly during external crises.

Zheng et al. (2022) examined the influence of employment across various industries on regional GDP within China's diverse regions, utilizing 2019 data. Their research employed multiple linear regression and cluster analysis methods. The study highlighted the significant impact of employment in five key sectors: wholesale and retail, manufacturing, construction, residential services, repair, and other

services; accommodation and catering; and leasing and business services on regional GDP. Cluster analysis revealed that coastal regions typically outperformed inland areas concerning the employed workforce and regional GDP. Notably, manufacturing employment was found to boost regional GDP by an average of 0.254 units, assuming that the other variables remained constant.

In recent years, researchers have expanded their analysis beyond traditional regression models and ventured into the realm of machine learning models to gain deeper insights into GDP analysis. Numerous studies have shown that neural network models outperform econometric models in terms of predictive accuracy.

For instance, Tkacz (2001) demonstrated that the use of leading neural network models significantly improved the precision of financial and monetary forecasts for Canada's output growth when compared to conventional linear and univariate time-series models. The neural network models yielded lower forecast errors for the annual growth rate of Canada's GDP, although the forecast improvements were relatively less pronounced when predicting quarterly GDP growth. Further supporting the superiority of machine learning techniques,

Jahn (2018) found that the artificial neural network (ANN) regression model exhibited remarkable capabilities in producing highly accurate predictions of GDP growth rates for 15 industrialized nations over the period from 1996 to 2016, surpassing the performance of conventional linear models.

Priambodo et al. (2019) opted for the K-Nearest Neighbour (K-NN) method over other prediction algorithms like neural networks and linear regression due to its simplicity and efficiency with small datasets. Their analysis involved a dataset containing commodity prices, exchange rates, government consumption, and export values spanning from 1980 to 2002. The K-NN regression predictions were compared with backpropagation neural network and multiple linear regression predictions, revealing that the K-NN regression method outperformed both the neural network and multiple linear regression in accurately forecasting Indonesia's GDP using a small dataset.

Likewise, Jena et al. (2020) highlighted the effectiveness of ANN models compared to traditional statistical methods in predicting outcomes with greater accuracy. By utilizing a Multilayer Artificial Neural Network (MLANN) model based on Q2-2020 data, they successfully forecasted the GDP of eight major countries, including the United States, Mexico, Germany, Italy, Spain, France, India, and Japan. The results demonstrated the MLANN model's capability in accurately predicting GDP, with the Mean Absolute Percentage Error (MAPE) for each country being less than 2%.

Jönsson (2020) investigated the business tendency survey data alongside the K-Nearest Neighbour algorithm to forecast Swedish GDP growth. The K-NN method utilized cross-validation to determine the optimal value of K. The findings, based on forecast errors using the Mean Absolute Error (MAE) and the Mean Squared Error (MSE), demonstrated that the K-NN method performed at least as well as the linear models in forecasting Swedish GDP growth.

Nadwah (2021) explored the application of the Weighted K-Nearest Neighbour (WK-NN) method to forecast the GDP of Indonesia's forestry and logging sector from Q1-2010 to Q4-2019. The GDP data was divided into train data and test data. The train data utilized the Multiple Input Multiple Output (MIMO) technique, while the test data utilized the most recent data pattern from the forestry and logging GDP dataset. The MIMO approach involved a single simulation and a specialized model to establish a forecast horizon. The performance of the models was assessed using the Mean Absolute Percentage Error (MAPE), and the value of K was determined based on the smallest MAPE.

Rigopoulos (2022) forecasted Greece's GDP and demonstrated that the application of the K-Nearest Neighbors (K-NN) algorithm for time series forecasting yields Mean Absolute Percentage Error (MAPE) results closely matching those of the well-established ARIMA (1,2,1) model. Both methods exhibited a strong alignment between their forecasted values and the actual values. Despite some

potential for improvement, the K-NN model displayed commendable accuracy for one-step-ahead forecasts, affirming its potential and promising forecasting capabilities.

These findings collectively highlight the growing significance of machine learning models in GDP analysis and their potential to enhance forecasting accuracy.

3. Methodology

3.1 Multiple Linear Regression (MLR)

Multiple linear regression is a statistical technique used in predictive modelling and statistical analysis to establish a relationship between a dependent variable and two or more independent variables. This method allows us to calculate the model's variation and the relative contribution of each independent variable to the total variance (Tranmer et al., 2020). The MLR model can be written in the form

$$y_i = \beta_0 + \sum_{j=1}^k \beta_j X_{ij} + \epsilon_i \quad (1)$$

for $i = 1, 2, \dots, n$, and $\epsilon_i \sim N(0, \sigma^2)$, where

y_i is the dependent variable of the i -th observation (the one being predicted),

β_0 is the regression intercept (the value of y when all predictors are 0),

β_j is the j -th predictor's regression slope,

X_{ij} the j -th predictor for the i -th observation,

ϵ_i represents the Gaussian error term for the i -th observation, and these errors are statistically uncorrelated.

The following steps are commonly used for performing a regression analysis:

1. **Data Pre-processing:** Clean and pre-process the data as needed. Handle missing values, deal with outliers, and perform any necessary transformations or scaling of variables.
2. **Data Plotting:** This step involves visualizing the relationship between the dependent variable and each independent variable to check for linearity using a scatterplot matrix. It's essential to ensure that the relationship is approximately linear before proceeding with regression analysis.
3. **Model Building:** This involves fitting a multiple linear regression model Eq. (1) using the chosen independent variables and estimating the regression coefficients (β 's) that best represent the relationship between the dependent variable and the independent variables.
4. **Model Evaluation:** After fitting the regression model, it is typically involves assessing the quality of the model using evaluation metrics such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), R -squared (R^2), adjusted R -squared, F -statistic, or other appropriate measures to determine the overall goodness-of-fit.
5. **Feature Selection:** If any independent variables are found to be statistically insignificant (p -values greater than a chosen significance level, e.g., 0.05), they can be considered for removal from the model. This step helps to simplify the model and identify the most relevant independent variables that significantly contribute to the prediction.
6. **Model Improvement:** Iterate through the process, refining the model as needed by making adjustment, incorporating domain knowledge, or adding new features to improve its performance. The refined model is then evaluated using the various evaluation metrics discussed in step (4).
7. **Final Model Deployment:** Once the model's performance is satisfactory, it can be deployed to make predictions on new, unseen data.

It's important to note that building a multiple linear regression model also involves considering assumptions such as linearity, independence of errors, constant variance (homoscedasticity), and

normality of errors. Checking these assumptions is a crucial part of the process and may require additional diagnostics and transformations. Some commonly used plots and tests for model diagnostics checking include the Residuals vs Fitted plot, Scale-Location Plot, Breusch-Pagan test for heteroscedasticity, Shapiro-Wilk test for normality, Variance Inflation Factor (VIF) for detecting multicollinearity, and others.

3.2 K-Nearest Neighbours (K-NN)

The K-Nearest Neighbours (K-NN) is a simple yet powerful machine learning algorithm. It is a supervised and non-parametric learning classifier that does not make any assumptions about the data being examined. While K-NN can be used for both regression and classification, we will focus on regression problems and construct an optimal regression model. In the regression task, K-NN takes the average (or sometimes the median) of the target values of its k nearest neighbours. The new data point is then predicted with this average value as the output.

In this analysis, we use the Euclidean distance method, which is the most widely used distance metric in K-NN. It is favoured for its simplicity and effectiveness when dealing with dimensions on the same scale. The Euclidean distance is commonly applied to calculate the distance between two real-valued vectors. It measures the length of a straight line connecting two query points and is expressed by the formula:

$$d(x, y) = \sqrt{\sum_{i=1}^n (y_i - x_i)^2} \quad (2)$$

However, it's important to note that in some situations, alternative distance metrics, such as the Manhattan distance, may be more appropriate. This is especially relevant when dealing with data that has varying scales or when certain features are more influential than others in determining similarity.

The K-Nearest Neighbours (K-NN) algorithm involves the following essential steps:

1. **Choose the Optimal k :** k is a hyperparameter that requires predefinition to make accurate predictions. Determining the number of nearest neighbours (k) significantly impacts the algorithm's performance. Techniques like cross-validation or other optimization methods can be used to find the optimal value of k .
2. **Select a Distance Metric:** To measure the similarity between data points, a suitable distance metric is chosen. Commonly used metrics include Euclidean distance Eq. (2) and Manhattan distance. However, depending on the data and problem, alternative distance metrics can be considered.
3. **Data Pre-processing:** Data Pre-processing: Prior to applying K-NN, the data is pre-processed by handling missing values, managing outliers, and performing necessary transformations or scaling of variables. Since K-NN relies on distance calculations, ensuring that all features are on the same scale is essential.
4. **Model Training:** K-NN is an instance-based algorithm, it stores the entire training dataset in memory for making predictions. During the training phase, the algorithm memorizes the training data to be used for subsequent predictions.
5. **Prediction:** For each new data point in the test dataset, K-NN identifies the k nearest neighbours from the training dataset based on the chosen distance metric.
6. **Averaging (Regression):** In regression tasks, the algorithm calculates the average (or median) of the target values of the k nearest neighbours to predict the output value for the new data point.
7. **Model Evaluation:** To assess the performance of the K-NN model for regression tasks, appropriate evaluation metrics such as Root Mean Squared Error (RMSE) and Mean Absolute

Error (MAE) are used to provide valuable insights into the model's accuracy.

8. **Tuning k :** The process of tuning k involves iterating through different values of k and evaluating the model's performance each time. The value of k that demonstrates the best performance on the validation data is selected as the optimal k for the model.

By following the above steps (1) to (8) outlined above, the K-NN algorithm can be effectively applied to various datasets, and its predictive capabilities can be accurately evaluated.

3.3 Artificial Neural Network (ANN)

Artificial Neural Network (ANN) is a computer system that imitates the distributed and parallel computing capabilities of the brain to perform complex tasks such as control and recognition (ada, 1995). Due to its remarkable ability to learn from data, ANN has gained widespread popularity in the field of machine learning. In our research, we use ANN method to tackle the prediction of Indonesia's Gross Domestic Product (GDP), leveraging its capacity to effectively handle complex and non-linear patterns in economic data. Figure 1 shows a simple architecture of an ANN model.

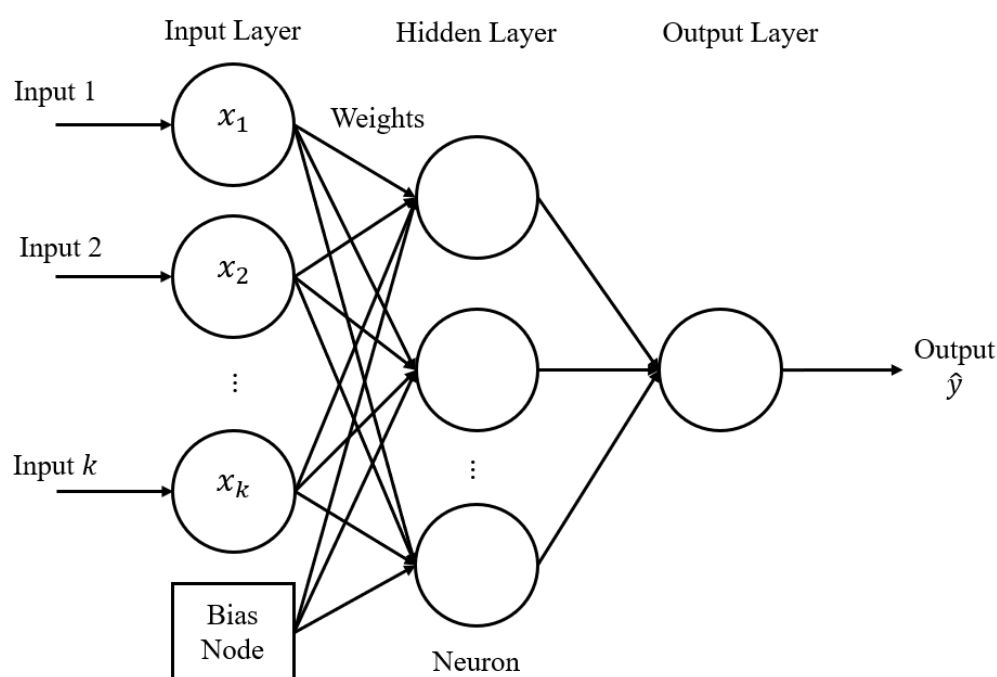


Figure 1. A simple diagram of an Artificial Neural Network (ANN).

The components of an ANN model are as follows, and they are illustrated in Figures 1 and 2.

- (a) **Neuron:** In an ANN, a neuron, also known as a node or unit, is a fundamental computational element that receives inputs, processes them through weighted connections, and produces an output using an activation function. Neurons play a key role in learning and modelling complex patterns in the data.
- (b) **Input Layer:** The input layer collects features for forecasting, and the number of neurons in this layer is determined by the number of input features in the dataset. Each neuron represents an input feature, and collectively they form the basis for data representation.
- (c) **Hidden Layer:** The hidden layer is situated between the input and output layers of an ANN and plays a crucial role in information processing. It processes the input data through a series of mathematical operations, extracting and capturing meaningful patterns from the input features.

- (d) **Output Layer:** The output layer produces predictions by aggregating the weighted sum of the neuron's inputs and adding the bias term. An activation function is then applied to determine the final output of a neuron, effectively translating the network's computations into predictions.
- (e) **Activation Function:** Activation functions introduce non-linearity to the network, allowing it to capture and learn complex data patterns. These functions serve as critical elements in enabling ANNs to model a wide range of problems with diverse and non-linear relationships. Some of the most commonly used activation functions in ANNs include: Sigmoid, Hyperbolic Tangent (Tanh), Rectified Linear Unit (ReLU), Leaky Rectified Linear Unit (Leaky ReLU), Exponential Linear Unit (ELU), Softmax, and others. The choice of activation function depends on the problem at hand, the network architecture, and the characteristics of the data.
- (f) **Weights and Biases:** Weights and biases are parameters that the network learns during training. Weights control the strength of connections between neurons in the network, while biases ensure that the network can learn non-linear relationships between inputs and outputs. The network adjusts the weights and biases during training to minimize the loss function.
- (g) **Loss Function:** The loss function measures the discrepancy between the predicted and actual output for a given input. In simple terms, the loss function is given by

$$L = (y_i - \hat{y}_i)^2 \quad (3)$$

Throughout the training process, the objective is to minimize the loss function Eq. (3), ensuring that the model becomes increasingly accurate in its predictions.

- (h) **Optimization Algorithm:** The optimization algorithm is responsible for updating the weights and biases in the network. The objective is to iteratively minimize the loss function, leading to a well-performing and predictive ANN model. ANNs utilize a variety of optimization algorithms, such as the Gradient Descent algorithm, Adam optimizer, Stochastic Gradient Descent (SGD), Mini-batch Gradient Descent, and others.

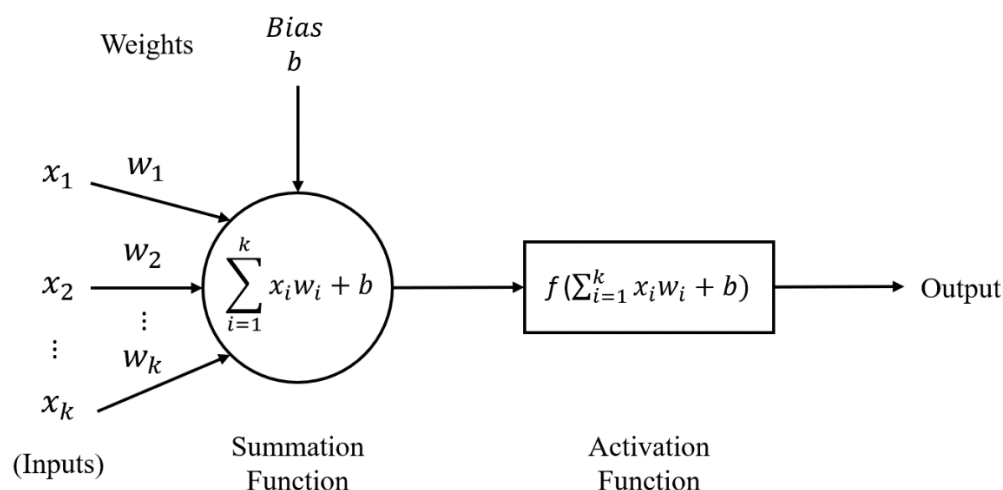


Figure 2. The mathematical model of a neuron in ANN.

The steps to construct an ANN model are as follows:

1. Import the required libraries and load the GDP dataset.
2. Split the dataset into input features and the target variable.
3. Further divide the dataset into training and test sets for model evaluation.
4. Define a function to create the Keras model and create a KerasRegressor object (for regression tasks) or a KerasClassifier object (for classification tasks).
5. Specify the hyperparameters to explore during the model tuning process. These may include the number of hidden layers, neurons in each layer, learning rate, activation functions, etc.

6. Utilize grid search or other hyperparameter optimization techniques to find the best combination of hyperparameters that result in the optimal model performance. Perform this search using the training data.
7. Create the final ANN model with the determined optimal number of neurons and other hyperparameters.
8. Compile the model by specifying the loss function, optimizer, and any additional evaluation metrics required.
9. Train the model using the training data, adjusting the model's weights and biases based on the training examples to minimize the loss function and improve performance.
10. Evaluate the trained model on the test set to assess its generalization ability and performance on unseen data. Utilize appropriate evaluation metrics for the specific task, such as Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) for regression or accuracy, precision, and recall for classification.

By following the steps (1) to (10) above, an ANN model can be effectively constructed and its predictive capabilities on the dataset can be evaluated.

3.4 Performance Metric

(a) Root Mean Square Error (RMSE)

Root Mean Square Error (RMSE) is a widely used evaluation metric for measuring the accuracy of forecasting models by comparing predicted and observed values. It quantifies the scatter of data around the line of best fit and is preferred over the normal mean square error (MSE) because its measurement range aligns with the analyzed data. The formula for RMSE is as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}$$

where

y_i is the i -th observed value,

\hat{y}_i is the i -th predicted value,

n is the number of observations.

(b) Mean Absolute Error (MAE)

Mean Absolute Error (MAE) is another frequently used performance metric for assessing the accuracy of forecasting models. It measures the average absolute error between the predicted results and the actual values. The MAE can be calculated by summing the absolute differences between each observation and its corresponding prediction, and then dividing this sum by the number of observations. The formula for MAE is given by:

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n}$$

4. Analysis and Results

4.1 Data Descriptive Statistics

In this project, we utilized a dataset obtained from the website of the Central Agency of Statistics, or Badan Pusat Statistik (BPS). The dataset provides a quarterly breakdown of Indonesia's Gross Regional

Domestic Product (GRDP) for its 38 provinces from 2010 to 2022, with values presented in hundreds of trillions of Rupiah (IDR). By aggregating all provincial GDP values, we were able to calculate Indonesia's national GDP, as illustrated in Figure 3. The plot in Figure 3 shows a positive trend in Indonesia's GDP. Although the economy has been severely affected by the pandemic in 2020, Indonesia has been able to quickly recover and continue to grow.

We then examine the four primary economic indicators that contribute to Indonesia's GDP: Net Export of Oil and Non-Oil Commodities (NE), Household Consumption (HC), Government Spending (GS), and Inventory Changes (IC). Figure 4 depicts the graphical summaries of these four explanatory variables.



Figure 3. Indonesia's 13-year GDP from Q1-2010 to Q4-2022.

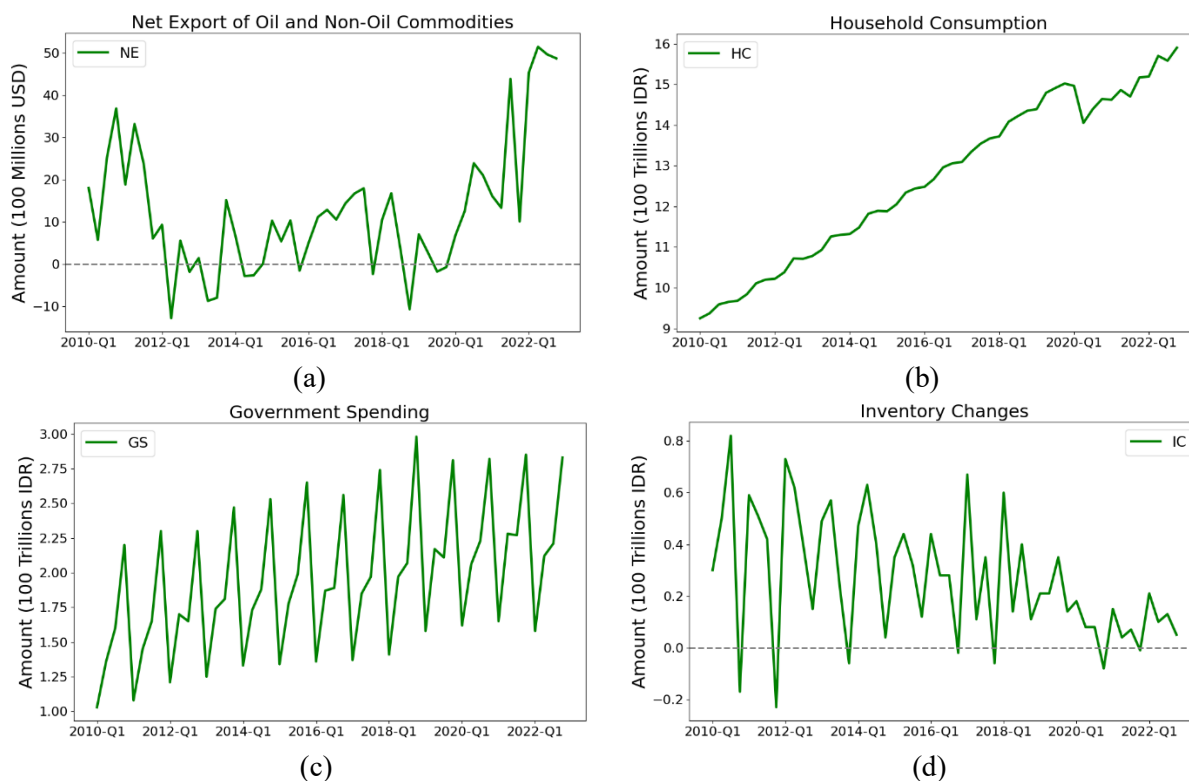


Figure 4. The four key economic indicators that contribute to Indonesia's GDP (BPS, 2022).

(a) Net Export of Oil and Non-Oil Commodities (NE)

Indonesia's net export was determined by deducting the value of imports from the value of exports, using two supporting datasets from the BPS website: 'Value of Export Oil&Gas – Non Oil&Gas (Million US\$), 2022', and 'Value of Import Oil&Gas - Non Oil&Gas (Million US\$), 2022'. These values were adjusted to hundreds of millions of USD.

Figure 4(a) shows that Indonesia's total export and import prices have experienced fluctuations but have recently surged due to high commodity prices and global demand. The country has a trade deficit, with the country importing more than exporting during certain periods of time.

(b) Household Consumption (HC)

Household consumption is an important variable included in the dataset as Indonesia's GDP. As seen in Figure 4(b), it follows a consistent positive increasing linear trend over the years, like Indonesia's GDP. This similarity is primarily due to household consumption expenditure being one of the largest contributors to total GDP. Although the pandemic had a significant impact on Indonesian spending power, encouragingly, it has shown a gradual recovery over time.

(c) Government Spending (GS)

Government spending is another significant variable included in Indonesia's GDP dataset. Figure 4(c) illustrates the cyclical pattern of the Indonesian government's expenditure, exhibiting an increasing trend over time. Notably, government spending consistently reaches its peak during the fourth quarter of each year and declines to its lowest point in the first quarter of the subsequent year. This is a result of the fiscal cycle's dynamics as the government needs to utilize budgets before the fiscal year closes, preventing unutilized funds from returning to the state treasury. Achieving optimal expenditure is crucial for fiscal goals, but transparency and efficiency are key to addressing concerns and maximizing budget impact (Irianto, 2021).

(d) Inventory Changes (IC)

Indonesia's GDP dataset also includes inventory changes as a variable. Figure 4(d) demonstrates that the Indonesian government's inventory changes follow some patterns. The final quarter of every year is when the government typically releases a large percentage of its inventories. There is also evidence of a slowdown in these changes over time.

4.2 Results

We split our dataset from Q1-2010 to Q4-2022 into 80-20 train-test sets. As there are 52 data points, 42 data points will be used for training, and 10 data points are used for testing. This split ensures the balance between training the model on sufficient data to learn patterns and testing its generalization to new data. This practice minimizes the risk of underfitting and overfitting, leading to a reliable model evaluation (Hastie et al., 2009).

Using the multiple linear regression (MLR) method, we compare the predicted and actual values for Indonesia's GDP, as depicted in Figure 5. The plot exhibits a close alignment between the predicted and actual values, providing compelling evidence of the predictive MLR model's accuracy.

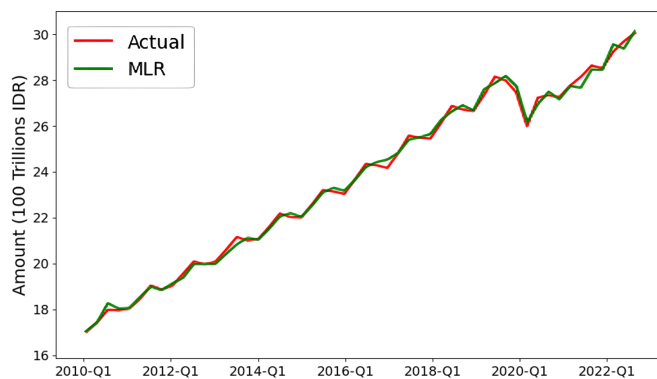


Figure 5. Predicted vs actual plot of the MLR model.

To validate the reliability of our regression findings, we conduct comprehensive model diagnostic assessments, encompassing residual plot, the Shapiro-Wilk normality test, the Breusch-Pagan test, and the Variance Inflation Factor (VIF). As illustrated in Figure 6, Table 1, Table 2 and Table 3, all results confirm that there are no significant issues with the assumptions underlying the MLR model.

The residual plot in Figure 6 shows that there is a slight degree of heteroscedasticity as seen from the increasing spread of residuals along the *x*-axis. However, linearity appears to be well held, as all the points seem randomly scattered.

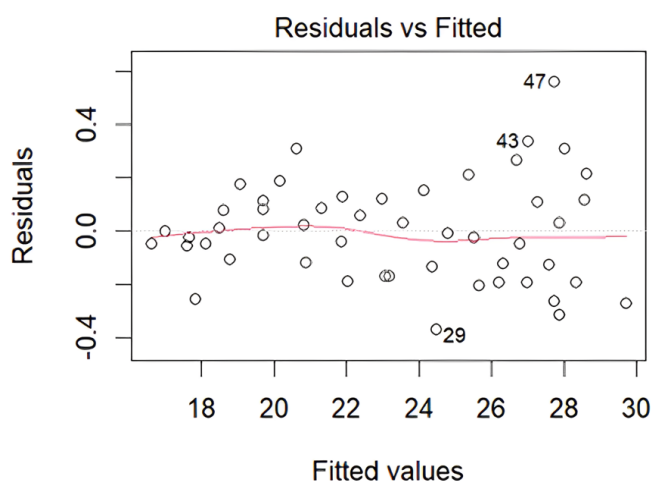


Figure 6. The residual plot of the MLR model.

Table 1. The Shapiro-Wilk normality test

| Shapiro-Wilk Normality Test | |
|-----------------------------|-----------------|
| <i>W</i> | <i>p</i> -value |
| 0.94893 | 0.02624 |

The results of the Shapiro-Wilk test, displayed in Table 1, show a calculated test statistic (*W*) of 0.94893, with a corresponding *p*-value of 0.02624. This indicates that the investigated MLR model adheres to the assumption of normality at the 1% significance level.

Table 2. The studentized Breusch-Pagan test.

| Studentized Breusch-Pagan Test | | |
|--------------------------------|-----------|-----------------|
| <i>BP</i> | <i>df</i> | <i>p</i> -value |
| 11.608 | 4 | 0.02051 |

In Table 2, the studentized Breusch-Pagan test yields a test statistic of 11.608, with a corresponding p -value of 0.02051. This indicates a potential presence of heteroskedasticity at the 5% significance level. Nevertheless, as the p -value exceeds 0.01, we can justify the assumption of homoscedasticity at the 1% level.

Table 3. The Variance Inflation Factor (VIF).

| Variables | VIF Value |
|-----------|-----------|
| NE | 1.20887 |
| GS | 2.695097 |
| IC | 2.451542 |
| HC | 1.488684 |

The VIF analysis presented in Table 3 shows that all predictors have VIF values below 3. Therefore, no additional investigation is necessary, and there are no significant concerns regarding multicollinearity among the predictors.

The high R^2 value of the model (99.76%) indicates that the four key economic indicators: Net Export of Oil and Non-Oil Commodities (NE), Household Consumption Expenditure (HC), Government Spending (GS), and Inventory Changes (IC), explain a significant large proportion of the variation in Indonesia's GDP. Thus, the predictive MLR model appears highly reliable in forecasting Indonesia's GDP.

In the K-NN method, we first determine the optimal value of k using the Mean Absolute Error (MAE) metric. Lower values of k indicate higher accuracy in the prediction. As seen in Table 4 and Figure 7, the optimal value is $k = 1$. This implies that using only the nearest data point to the target point (i.e., $k = 1$) provides optimal results. This is likely due to the potential over-generalization and decreased accuracy that may occur when considering a larger number of data points with higher k values.

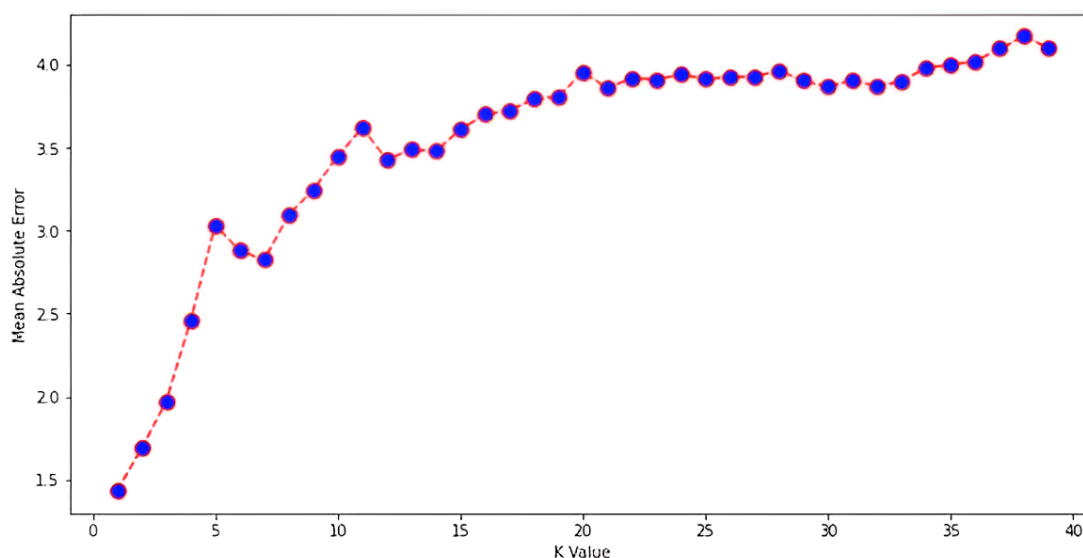
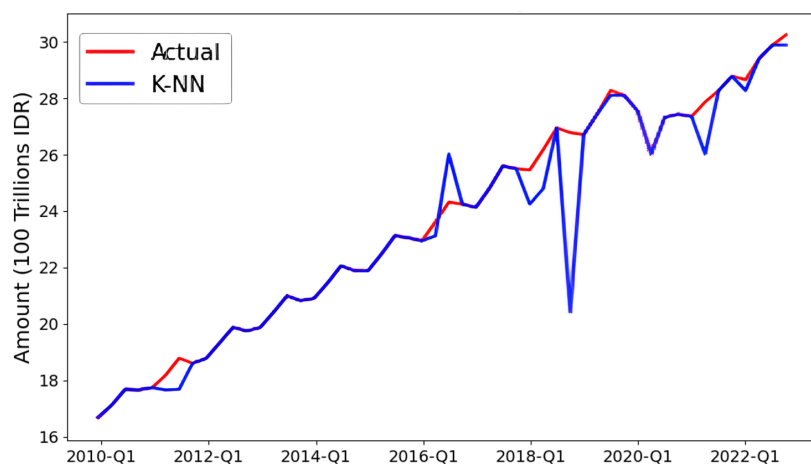


Figure 7. Plot of MAE values for k from 1 to 39.

Table 4. The MAE for the first 10 k values in the K-NN method.

| k | Mean Absolute Error (MAE) |
|-----|---------------------------|
| 1 | 1.435 |
| 2 | 1.689 |
| 3 | 1.972 |
| 4 | 2.457 |
| 5 | 3.034 |
| 6 | 2.879 |
| 7 | 2.824 |
| 8 | 3.096 |
| 9 | 3.245 |
| 10 | 3.448 |

Using the optimal value of $k = 1$, we compared the predicted and actual values for Indonesia's GDP using the K-NN regression model, as shown in Figure 8. While the K-NN method successfully predicts most of the GDP values that align well with the data trend, there are some notable anomalous points that exhibit significant deviations from the actual GDP values.

Figure 8. Predicted vs actual plot of the K-NN regression model with $k = 1$.

In the Artificial Neural Network (ANN), we employ three dense layers, with the first two utilizing the Rectified Linear Unit (ReLU) activation function, and the last layer employing a linear activation function. The choice of ReLU is preferred due to its computational efficiency and its tendency to show better convergence performance than sigmoid and tanh neurons (Krizhevsky et al., 2017). The linear activation function in the third layer is well-suited for predicting continuous values of the target variable, which aligns perfectly with the GRDP as the response variable in this study.

The predicted ANN model is optimized using the Adam optimizer, and we adopt the MAE as the loss function. The Adam optimizer converges faster than other optimization algorithms (Zaheer et al., 2019), while MAE is robust to outliers and provides an interpretable measure of error (Trevisan, 2022). This combination leads to better performance, faster convergence, and an accurate measure of the error in predictions. The optimal number of neurons in the first layer is determined through grid search cross-validation. We set the appropriate epochs, batch size, verbose mode, and validation set to prevent overfitting. Figure 9 shows the plot of the generated ANN predictions compared to the actual values of Indonesia's GDP from Q1-2020 to Q4-2022.

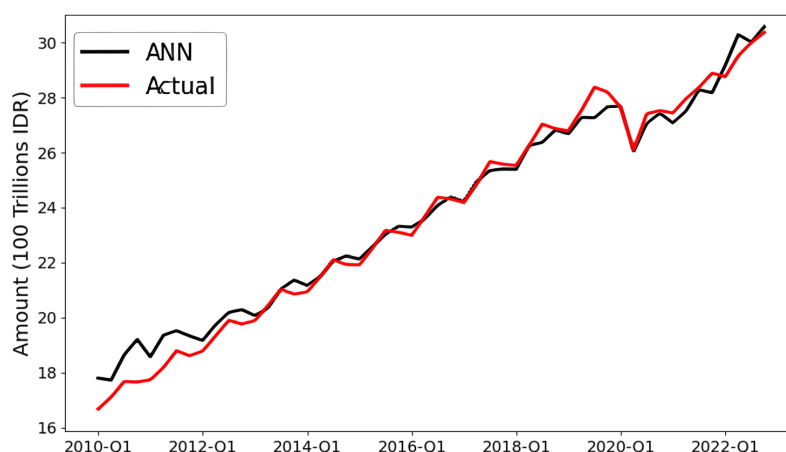


Figure 9. Predicted vs actual plot of the Artificial Neural Network (ANN) model.

The plot in Figure 9 demonstrates that the ANN model follows the actual GDP trend, with the predicted values being very close to the actual values. The plot indicates that the ANN model has performed well in predicting the GDP values, with a very small deviation from the actual values. This suggests that the ANN model can be used for accurate forecasting of Indonesia's national GDP values in the future, provided that the input variables remain unchanged.

Table 5. RMSE and MAE outputs of MLR, K-NN and ANN models.

| Output | MLR | K-NN | ANN |
|--------|--------|--------|--------|
| RMSE | 0.1485 | 2.2116 | 0.4204 |
| MAE | 0.1217 | 1.4345 | 0.5552 |

To determine the superior predictive model, we compared the performance of the MLR, K-NN, and ANN models using two fundamental metrics: RMSE and MAE. The values in Table 5 represent the models' performance during the testing phase, maintaining consistency with the earlier methodology. While the plots visualize both training and testing data, the RMSE and MAE values exclusively reflect performance during the testing phase for each model. The MLR model exhibited the lowest values for both RMSE and MAE, signifying its highest level of accuracy among the three models. The MLR model demonstrates an approximate 93.3% improvement over the K-NN model and an approximate 64.7% improvement over the ANN model based on the RMSE values. Based on these results, we conclude that the MLR model is the most suitable for predicting Indonesia's GDP in this study.

Table 6. Q4-2022 GDP Predicted vs. Actual values by MLR, K-NN and ANN models.

| Model | Q4-2022 GDP | Difference |
|--------|-------------|------------|
| Actual | 30.31 | – |
| MLR | 30.39 | 0.08 |
| K-NN | 29.94 | – 0.37 |
| ANN | 30.47 | 0.16 |

To further validate the MLR model's effectiveness, we forecasted Indonesia's Q4-2022 GDP final data point using historical data with MLR, K-NN, and ANN models. The comparison of predicted and actual GDP values for the same period is presented in Table 6. Table 6 shows that the MLR and ANN models slightly overestimate Indonesia's Q4-2022 GDP final data point, while K-NN underestimates it. The MLR model exhibits the highest accuracy with a prediction difference of only 0.08, making it the most favourable choice for predicting Indonesia's GDP.

5. Conclusion and Discussion

In conclusion, this study demonstrates that the MLR model is the most reliable approach for forecasting Indonesia's GDP, using four key economic indicators: Net Export of Oil and Non-Oil Commodities (NE), Household Consumption (HC), Government Spending (GS), and Inventory Changes (IC). The MLR model outperforms the machine learning models, K-NN and ANN, due to its ability to recognize significant variables, simplicity, interpretability, and lower risk of overfitting. The K-NN model is not preferred over MLR and ANN due to its larger dataset requirement, lack of interpretability, inability to prioritize significant features, and sensitivity to input feature scales. The ANN models' high complexity and time-consuming tuning make them more prone to overfitting, highlighting that ANN is not always necessary, and a simpler model like MLR suffices to provide more interpretable results.

However, this study has a limitation, as it solely utilizes four explanatory variables to predict Indonesia's GDP. While the MLR model is effective in GDP prediction, it may not encompass all economic factors that influence it. Hence, incorporating additional explanatory variables such as stock market indices, consumer price index, gross fixed capital formation, and employment rate, which are also key economic indicators, could be considered to enhance the model's comprehensiveness.

To further advance the research, investigating the impact of the COVID-19 pandemic on Indonesia's GDP and its implications for future forecasts would be valuable. Furthermore, extending the analysis to explore regional variations in GDP and their correlation with local economic factors could provide deeper insights.

For future work, the research methodology could be expanded to include time series models, such as ARIMA, VAR, and GARCH. These time series models capture various seasonal and cyclical patterns in economic data, potentially improving the accuracy of Indonesia's GDP forecasts.

In summary, the MLR model emerges as the most suitable choice for predicting Indonesia's GDP in this study, yet further exploration and augmentation of the research framework offer promising avenues for refining and enhancing the forecasting capabilities.

6. Acknowledgements.

I extend my sincerest gratitude to Dr. Soo Huei Ching and Mrs. Bernadette Lawardi for their invaluable guidance and support throughout the development of this manuscript. Their expert insights and encouragement were instrumental in ensuring its success. Furthermore, I would like to express my appreciation to the Editorial team of JOSMA for their thorough review of this work.

7. References

- Badan Pusat Statistik (2022). *Gross Regional Domestic Products (Expenditure)*. Retrieved from <https://www.bps.go.id/indicator/171/533/1/-seri-2010-2-pdrb-atas-dasar-harga-konstan-menurut-pengeluaran-2010-100-.html>.
- Badan Pusat Statistik (2022). *Value of Export Oil&Gas - Non Oil&Gas (Million US\$)*. Retrieved from <https://www.bps.go.id/indicator/8/1753/1/nilai-ekspor-migas-nonmigas.html>.
- Badan Pusat Statistik (2022). *Value of Import Oil&Gas - Non Oil&Gas (Million US\$)*. Retrieved from <https://www.bps.go.id/indicator/8/1754/13/nilai-impor-migas-nonmigas.html>.
- Callen, T. (2012). *Gross Domestic Product: An Economy's All*. *International Monetary Fund: Washington, DC, USA*.
- Irianto, K.B. (2021). Surat Opini: Mengapa Belanja Pemerintah Selalu Meningkatkan di Akhir Tahun? *CNBC Indonesia*. Retrieved from <https://www.cnbcindonesia.com/opini/20211223183110-14-301570/mengapa-belanja-pemerintah-selalu-meningkat-di-akhir-tahun>.

- Desai, N.A., & Bhatia, M. (2016). A Multiple Variable Regression Model for Gross Domestic Product Growth Rate Prediction in India Using Key Macroeconomic Indicators. *Journal of Economics and Finance*. 7(2): 47 – 51.
- Dudu, L., & Moscu, R.G. (2016). The Practical Use of The Linear Multiple Regression in The Complex GDP Analysis. *Knowledge Horizons. Economics*. 8(2): 74 – 79.
- Hassoun, M.H. (1995). Fundamentals of Artificial Neural Networks. *A Bradford Book: The MIT Press, Cambridge, MA*. Retrieved from https://www.researchgate.net/publication/3078997_Fundamentals_of_Artificial_Neural_Networks-Book_Reviews.
- Hastie, T., Tibshirani, R., & Friedman, J.H. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer.
- Jahn, M. (2018). *Artificial Neural Network Regression Models: Predicting GDP growth* (No. 185). HWWI Research Paper.
- Jena, P.R., Majhi, R., Kalli, R., Managi, S., & Majhi, B. (2021). Impact of COVID-19 on GDP of Major Economies: Application of The Artificial Neural Network Forecaster. *Economic Analysis and Policy*. 69: 324 – 339.
- Jönsson, K. (2020). Machine Learning and Nowcasts of Swedish GDP. *Journal of Business Cycle Research*. 16(2): 123 – 134.
- Krizhevsky, A., Sutskever, I., & Hinton, G.E. (2017). ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*. 60(6): 84–90.
- Nadwah, A., Nusyirwan, N., & Faisol, A. (2021). Implementasi Weighted K-Nearest Neighbor Untuk Peramalan Data Deret Waktu. *Jurnal Riset dan Aplikasi Matematika*. 5(2): 111 – 117.
- Nwabueze, J.C. (2009). Causal Relationship Between Gross Domestic Product and Personal Consumption Expenditure of Nigeria. *African Journal of Mathematics and Computer Science Research*. 2(8): 179 – 183.
- Priambodo, B., Rahayu, S., Naf'an, E., Handriani, I., Putra, Z.P., Nseaf, A.K., Setiawan, D., & Jumaryadi, Y. (2019). Predicting GDP of Indonesia Using K-Nearest Neighbour Regression. *In Journal of Physics: Conference Series* (Vol. 1339, No. 1, p. 012040). IOP Publishing.
- Rigopoulos, G. (2022). Univariate Time Series Forecasting Using K-Nearest Neighbors Algorithm: A Case for GDP. *International Journal of Scientific Research and Management (IJSRM)*. 10(9): 3807-3815.
- Stanić, S., & Račić, Ž.V. (2019). Analysis of Macroeconomic Factors Effect to Gross Domestic Product of Bosnia and Herzegovina Using the Multiple Linear Regression Model. *Economics*. 7(2): 91 – 97.
- Tkacz, G. (2001). Neural Network Forecasting of Canadian GDP Growth. *International Journal of Forecasting*. 17(1): 57 – 69.
- Tranmer, M., Murphy, J., Elliot, M., & Pampaka, M. (2020). Multiple Linear Regression (2nd Edition). *Cathie Marsh Institute Working Paper 2020-01*. Retrieved from <https://hummedia.manchester.ac.uk/institutes/cmist/archive-publications/working-papers/2020/2020-1-multiple-linear-regression.pdf>.
- Trevisan, V. (2022). Comparing Robustness of MAE, MSE, and RMSE. *Towards Data Science*. Retrieved from <https://towardsdatascience.com/comparing-robustness-of-mae-mse-and-rmse-6d69da870828>
- Zaheer, R., & Shaziya, H. (2019). A Study of the Optimization Algorithms in Deep Learning. *In Proceedings of the Third International Conference on Inventive Systems and Control (ICISC)*. Coimbatore, India. (pp. 536-539).
- Zheng, W., Xu, Y., Yang, J., & Yang, S. (2022). Research on the Impact of Employment on GDP Based on Multiple Linear Regression Model. *Economic Management Journal*. 11(1).