Accredited Ranking SINTA 2
Decree of the Director General of Higher Education, Research, and Technology, No. 158/E/KPT/2021
Validity period from Volume 5 Number 2 of 2021 to Volume 10 Number 1 of 2026

Published online on: **http://jurnal.iaii.or.id**

JURNAL RESTI
(Rekayasa Sistem dan Teknologi Informasi)
Vol. 6 No. 6 (2022) 1057-1063 ISSN Media Electronic: 2580-0760

Indonesian Crude Oil Price (ICP) Prediction Using Multiple Linear Regression Algorithm

Des Suryani[1], Mutia Fadhilla[2], Ause Labellapansa[3]
[1,2,3]Informatics, Engineering, University of Islam Riau
[1]des.suryani@eng.uir.ac.id, [2]tiafadhilla@eng.uir.ac.id, [3]ause.labella@eng.uir.ac.id

Abstract

Crude oil prices play a significant role in the global economy, therefore accurate prediction of oil prices is very important. Therefore, a forecasting model is needed to predict Crude Oil Prices. The purpose of this study is to forecast the price of crude oil from Indonesia (ICP). The data source is from a website published by the Ministry of Energy and Mineral Resources (ESDM), namely monthly crude oil price data specifically for six main types of crude oil: SLC, Attaka, Duri, Belida, Banyu and SC. The data used is data for a period of 5 years (2018 – 2022). The data available is in the form of time series data. Dated Brent combined with the Alpha factor for each month and year is a reference in determining the ICP price. Forecasting Indonesian crude oil prices in the future is based on the historical oil price of the previous period. The Data Mining algorithm used for forecasting is Multiple Linear Regression. The dataset processed using training data is 80%, and testing data is 20%. The model produced, on average, has a good level of accuracy in calculating MAPE where for SLC = 9%, Attaka = 45%, Duri = 126%, Belida = 33%, Banyu = 150% and SC = 50%. Based on the MAPE calculation value, the Linear Regression Equation to predict Indonesian Crude Oil Prices (ICP) shows that the model produced by SLC crude oil is very good. Attaka, Belida and SC crude oil yielded fair yields and Duri and Banyu crude oil yielded poor yields.

Keywords: crude oil, ICP, time series, prediction, linear regression

1. Introduction

In the world economy, crude oil plays an important role, therefore accurate prediction of oil prices is very important[1],[2]. Crude oil price fluctuations and the accurate predictions are a concern for many industry practitioners, researchers and policy makers[3],[4]. One of the plantation products that makes up Indonesia's primary export is palm oil, often known as crude palm oil. In comparison to OPEC nations, Indonesia tends to import oil or has significant local oil demands. Indonesian crude oil is not very competitive or tends to be imported by other nations [5],[6].The market prospects for palm oil have increased significantly from year to year. Price is an important factor in determining the selling value of the resulting product. Prices affect producer profits. Price is also decisive for consumers to buy, so it is very important to monitor and predict the price of crude oil or Indonesian Crude Oil Price (ICP). Crude oil is one of the main resources in the energy sector, and the efficiency of this industry partly depends on the price of this resource. Industries, governments, and people can all benefit greatly from these pricing estimates [7]. Therefore, it is important to buy it at minimum price, which requires tools for market price forecasting. As well it can be useful in Strategic Trade Theory that can be improved by predictions[8], [9]. Forecasting crude oil prices is very important to help businesses and governments make decisions about the energy market and can reduce the impact of price fluctuations. In order to accomplish this, we require a program that can assist in estimating the price of palm oil, particularly for the domestic market.

Several studies have been conducted to predict the Indonesian Crude Oil Price (ICP) price pattern. Muhammad Hussein (2020) indicated that oil prices would be influenced by two factors, which are demand and income[10][11]. Multiple linear regression was used in Ahmad Fitri Boy's research (2020) to forecast the price of Indonesian Crude Oil (ICP) in the domestic market [12]. This research was conducted 1) to identify the characteristics of crude oil price data in Indonesia, 2) to produce forecasting models with data mining algorithms that are suitable for volatile crude oil prices, 3) the research results can be used in a decision making for stakeholders and 4) add knowledge of machine learning in the field of data mining science.

Accepted: 14-10-2022 | Received in revised: 08-12-2022 | Published: 30-12-2022

The urgency of this research is to predict the ICP price for the future period based on the ICP price in the previous period, where the available data is in the form of times series data. Dated Brent combined with the Alpha factor is a reference in determining the ICP price. Forecasting Indonesian crude oil prices in the future is based on the historical oil price of the previous period with data mining algorithms. The Data Mining Algorithm used for forecasting is Multiple Linear Regression. Identification of relationships and their impacts on object values is the process of regression. Finding a function that best represents the data by reducing the error or difference between the predicted and actual values is the goal of regression analysis [13]. Multiple Linear Regression is a method to study the relationship between variables in the data forecasting process. Therefore, the Multiple Linear Regression method is the right algorithm for forecasting problems. Dated-Brent combined with the Alpha factor is a reference in determining the ICP price. The ICP formula consists of Dated-Brent plus Alpha. Forecasting the price of Indonesian crude oil in the coming period is based on historical oil prices in the previous period. The Multi Linear Regression method used in this study is to predict Dated-Brent and Alpha values. The results of Dated-Brent and Alpha predictions indicate the ICP obtained from the sum of the Dated-Brent prediction values and Alpha prediction values. Performance metrics that will be used are Mean Square Error (MSE), Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). A number of predicting accuracy metrics including Mean Square Error (MSE), Root Mean Square Error (RMSE).

2. Research Methods

2.1 Research Stage

In data mining, information is stored electronically and is searched automatically using computer algorithms. Data mining is the process of solving problems by analyzing data that is already available in databases. A customer database with customer profiles could help solve this problem [9], [14]. The research method is a sequential process in research activities. In this study, several research steps will be carried out, as depicted in Figure 1.

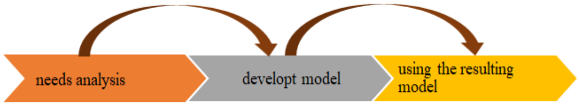

Figure 1. Research Stage

Analyzing the needs that consist of : a). Define goals and plan, b). Collecting data. The Ministry of Energy and Mineral Resources website contains five years' worth of Indonesian crude oil price data, which serves as the data source (2018-2022), c). Reviewing data Developing the Model that include a). Data Preparation. b). Model Testing. Data preparation is a very important step for predictive model design. Data preprocessing is a process of cleaning, transforming, and reducing data before analysis. Because real-time database data is frequently incomplete and inconsistent, preprocessed data is frequently used. This can result in inaccurate data mining results. As a result, the data preprocessing steps listed below must be completed in order to improve the quality of the data that will be analyzed: 1). Data Cleaning. The data collected may have many parts that do not fit, and some parts are missing, so a data cleaning process is required. 2). Data Transformation. Data transformation is used to transform data into a suitable form. 3). Data reduction. Data analysis using large data sets is difficult, so data reduction techniques are needed to improve storage efficiency and reduce data storage and analysis costs. The degree of deviation or error between the predicted and actual data is known as predictive accuracy [15].

In using the resulting model, validation is a very important step in modelling to see how reliable the model is for decision-making. Errors in making predictions are caused not only by factors causing errors but also by the inability of the prediction model to recognize other factors in the data set that affect the variance in predictions. Some ways to calculate error magnitude are MSE, RMSE and MAPE. MSE is the average of the squared difference between the predicted and observed values, RMSE is the root of MSE, and MAPE is the average of the absolute difference between the predicted and actual values[16]. Deployment, evaluate the quality and effectiveness of one or more models submitted in the modelling phase before placing them for use in the field.

2.2 *Data Analysis Method*

In this study, the first dataset test performed was the preprocessing process. Preprocessing data involves transforming raw data into a form that is more easily understood. This process is important because raw data often do not have a regular format. Furthermore, data mining also cannot process raw data, so this process is important to facilitate the next cycle, which is data analysis.

This process generates a new dataset. Then the modelling process is carried out using the Linear Regression algorithm, one of the prediction algorithms in the field of data mining. The next stage is the performance test process to see the accuracy of the predictions (Validation Stage). The last stage is to test the prediction of the testing data. The proposed design is shown in Figure 2.

DOI: https://doi.org/10.29207/resti.v6i6.4590

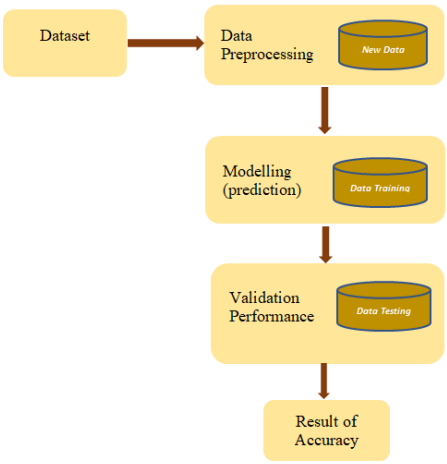

Figure 2. Data Analysis Method

2.3 *Linier Regression*

The linear regression model is a reliable tool for analyzing real-world data. There are many benefits to using linear regression, such as the fact that the linear regression model in training is faster than many predictive models [16],[17],[19]. Linear regression is a statistical method that is used to determine the strength of the relationship between a dependent variable and one or more independent variables. It can also be used to determine which independent variables are not related to the dependent variable, and which independent variables contain information that is redundant with information about the dependent variable. Linear regression models are easy to use and can be implemented quickly using memory resources [18],[19]. The statistical method known as regression analysis is used to examine the relationship that exists between two or more variables. Dependent (response) variables are typically measured on a scale, and one or more predictor variables are used to help explain or predict changes in the dependent variable(s). The relationship between these variables is modelled as a function (equation), for example, a linear function. The purpose of the regression model is to obtain the estimated parameters (coefficients) of the regression modelIn its simplest form, the independent variable (X) and dependent variable (Y) have the following equations:

$$Y = a + bX \quad (1)$$

Here, b denotes the direction or beta coefficient, while a represents the intercept. There is only one line equation function Y = a + bX that can be created from two points with dissimilar coordinates, namely (X_1, Y_1) and (X_2, Y_2).

Multiple linear regression analysis is an analysis that includes multiple independent variables. Determine whether there is a significant relationship between two or more independent variables (X_1, X_2, X_3, ..., X_k) and the dependent variable (Y) using the multiple linear regression technique. The following is an example of a population multiple linear regression model:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + \cdots + b_nX_n \quad (2)$$

Information:
Y: Dependent Variable, X (1,2,3,…): Independent Variable, a: Constant value, b (1,2,3,…): Regression coefficient value

The magnitude of the model prediction error is the error that occurs between the predicted data and the actual data. The error is represented using the original Mean Squared Error (MSE), which is the mean of the squared difference between the predicted and observed values; the original mean square error (RMSE) is the original of the MSE. RMSE is a good measure of accuracy, but just to compare the error different model predictions for variables certain and not between variables, because depending on the scale[13],[19], [20].

The formulas for MSE and RMSE are as follows:

$$\text{MSE} = \sum \frac{(Y' - Y)^2}{n} \quad (3)$$

$$\text{RMSE} = \sqrt{\sum \frac{(Y' - Y)^2}{n}} \quad (4)$$

Information:
Y'= Predicted value, Y = Actual value, n = Number of data

Mean Absolute Percentage Error (MAPE) is obtained by adding as a whole and subtracting the value of the actual data with the forecasting data. Then, this value is divided by the actual data (the absolute value is required), multiplied by 100, and divided by the number of existing data. The absolute value in this term is the value that remains positive even if it is negative.

The Mean Absolute Percentage Error (MAPE) formula is as follows:

$$\text{MAPE} = \frac{\sum_{t=1}^{n} \left| \left(\frac{A_t - F_t}{A_t} \right) 100 \right|}{n} \quad (5)$$

Information:
At = Actual demand to t, Ft = forecasting result to t
N = the amount of forecasting data, The MAPE formula's absolute symbol denotes that the computation result's negative value will remain positive.

2.4 *Dated Brent and Alpha*

The Ministry of Energy and Mineral Resources (ESDM) has issued a regulation regarding the Indonesian Crude Price (ICP) formula. This pricing no longer uses the Platts and RIM Intekegence formulas.

The government changed the formula for the Indonesian oil price (ICP) using the Dated Brent reference. The use of Dated Brent as a price basis

DOI: https://doi.org/10.29207/resti.v6i6.4590

provides the basic advantage that it is more independent; it is more difficult to be influenced or manipulated by certain parties because many countries use it as the benchmark price for Brent. The ICP formula consists of Dated Brent plus Alpha, calculated by considering the suitability of crude oil quality, developments in international crude oil prices and national energy security. Alpha will be determined monthly by the Minister of Energy and Mineral Resources.

3. Results and Discussions

3.1 *Data Set*

The data was obtained through a website published by the Ministry of Energy and Mineral Resources (ESDM). The data source used in this study is monthly crude oil price data for 55 crude oil over four years (2018-2022). Data for 2018-2022 is used in this study because the ICP formula consists of Dated-Brent plus Alpha according to the Ministry of Energy and Mineral Resources, which came into force in July 2016. However, the data available do not contain complete data for 2016-2017. Based on the data obtained from 2018 to 2022, the Dated-Brent and Alpha values are complete in each month of each year compared to 2016-2017 so in this study, data was used for 2018-2022. The 55 types of crude oil are Indonesia's main crude oil: SLC, Duri, Attaka, Belida, Banyu Urip and Senipah Condensate. Meanwhile, other Indonesian Crude Oils are Arjuna, Anoa, Arun Condensate, Bekapai, Belanak, Bontang Return Condensate (BRC), Bentayan, Bula, Bunyu, Cepu, Cinta, Camar, Geragai/Makmur, Geragai Condensate/Makmur Condensate, Handil Mix, Jambi, Jene/Pendopo, Jatibarang, Kaji/Matra, Grouper. In this study, the data used are the main crude oils, namely SLC, Attaka, Duri, Belida, Banyu and Senipah Condensate (SC). Table 1 shows the data set of SLC crude oil price.

Table 1. Data Set of SLC Crude Oil Price

Year	Month	Dated-Brent	Alpha	Price
2018	1	69.18	-3.35	65.83
2018	2	65.19	-2.88	62.31
2018	3	65.90	-3.05	62.85
2018	4	71.80	-3.41	68.39
…	…	…	…	…
2022	1	87.22	-1.35	85.87
2022	2	98.19	-2.26	95.93
2022	3	118.81	-5.76	113.05
2022	4	104.39	-0.82	108.57

3.2 *Modeling Result*

In modeling stage can do : (1) Predicting Indonesian Crude Oil Price (ICP) using Dated-Brent and Alpha calculations. Then the stages in the modelling process for ICP prediction are carried out with the following stages: a). Date-Brent Prediction using Multiple Linear Regression, b). Alpha Prediction using Multiple Linear Regression, c). ICP Price = Predicted Date-Brent + Predicted Alpha. (2) Implementation, the dataset consists of 51 actual data, 80% trained data and 20% testing data. The coding results are obtained through the Python programming language with a Linear Regression process from each of the six crude oil data: SLC, Attaka, Belida, Banyu, Duri and SC.

Date-Brent Prediction

Table 2. Results of Linear Regression for Variable Year, Month and Date-Brent

Type of Crude Oil	Intercept	X_1	X_2
SLC, Attaka, Belida, Banyu, Duri and SC	-10922.787586	5.44267	-0.23721

Based on the results in table 2, the Linear Regression Equation can be determined. Then the Dated-Brent prediction is made based on the Intercept, Year (X_1) and Month (X_2). The modelling for the Multiple Linear Regression algorithm in this study is Indonesia's Main Crude Oil, namely SLC, Attaka, Duri, Belida, Senipah Condensate (SC) and Banyu Urip. The Linear Equation Model for Date-Brent prediction.

Multiple Linear Equation Model for Major Crude Oil Date-Brent predictions :

$$Y = -10922.787585757089+(5.44267372* X_1) + (-0.23720747* X_2)$$

From the above equation to predict Dated-Brent crude oil from SLC, Attaka, Duri, Belida, Senipah Condensate (SC) and Banyu Urip in October – December 2022, the following results are obtained:

Table 3. Prediction Results of Date-Brent

Month	Prediction Date-Brent
October	79,9266
November	79,68939
December	79,45219

Figure 2 shows the Dated-Brent Value in 2022, where the Actual Value starts in January-September and the Predicted Value from October-December 2022.

Alpha Prediction

Results of Linear Regression for Alpha can be seen in table 4.

Table 4. Results of Linear Regression for Alpha

Type of Crude Oil	Intercept	X_1	X_2
SLC	382.52115517	-0.18885	-0.23206
Attaka	1458.9371354	-0.72241	-0.05507
Duri	-2663.638155	1.32005	-0.04473
Belida	694.13270916	-0.34371	-0.05601
Banyu	1135.5265597	-0.56029	-0.09717
SC	-1262.0086676	0.624306	-0.07675

DOI: https://doi.org/10.29207/resti.v6i6.4590

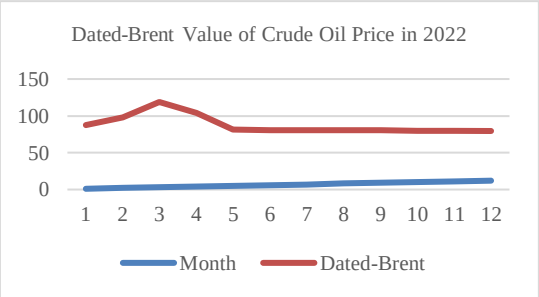

Figure 2. Dated-Brent Value of 2022

Linear Equation Model for Alpha prediction Multiple Linear Equation Model for SLC Crude Oil Alpha prediction:

$\hat{Y} = 382.521155 + (-0.188845*X_1) + (-0.23206 * X_2)$

Multiple Linear Equation Model for Attaka Crude Oil Alpha prediction:

$\hat{Y} = 1458.937135 + (-0.72244* X_1) + (-0.05507 * X_2)$

Multiple Linear Equation Model for Prediction of Thorny Crude Oil Alpha:

$\hat{Y} = -2663.6381549620346 + (1.32004921* X_1) + (-0.04472634 * X_2)$

Multiple Linear Equation Model for Belida Crude Oil Alpha prediction:

$\hat{Y} = 694.1327091550935 + (-0.34371446* X_1) + (-0.05600705 * X_2)$

Multiple Linear Equation Model for Banyu Crude Oil Alpha prediction:

$\hat{Y} = 1135.526559726688 + (-0.56028636* X_1) + (-0.09716756 * X_2)$

Multiple Linear Equation Model for SC Crude Oil Alpha prediction:

$\hat{Y} = -1262.008667673944 + (0.62430596* X_1) + (-0.07675289* X_2)$

The results are based on the above equation to predict the price of kerosene for Duri, SLC, Attaka, Senipah Condensate (SC), Belida, and Banyu Urip in October 2022, shown in table 5.

Table 5. Alpha Prediction Results

Type of Crude Oil	October	November	December
SLC	5,054139	5,009413	4,964687
Attaka	-1,418	-1,47401	-1,53001
Duri	1,655864	1,558697	1,461529
Belida	-0,42955	-0,5063	-0,58305
Banyu	-1,64451	-1,87657	-2,10863
SC	-2,3866	-2,4416	-2,4967

The graph in Figure 3 shows the Alpha value of one of the Main Crude Oils, which is Attaka, in 2022, where the actual value starts in January-September and the predicted value from October-December 2022.

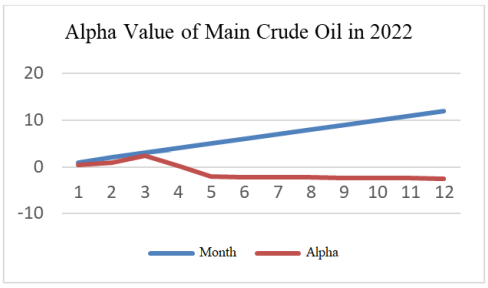

Figure 3. Alpha Value of Attaka Crude Oil in 2022

Figure 4 shows the Alpha values for the six main types of Crude Oil (SLC, Attaka, SC, Banyu, Belida and Duri) from January-December 2021.

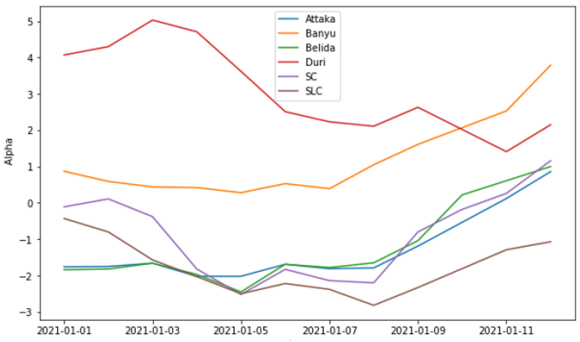

Figure 4. The alpha value of six types of Indonesian crude oil in 2022

ICP Price Prediction

The Indonesian Crude Price, especially the six main types of crude oil, is predicted by adding up the prediction results of Date-Brent and Alpha. For ICP predictions in October-December 2022, the prediction results are shown in Table 6.

Table 6. ICP Prediction Results October-December 2022

Month	Type of Crude Oil	Date-Brent Prediction Result	Alpha Prediction Result	ICP Prediction
October	SLC	79,9266	-1,64451	78,28209
	Attaka	79,9266	-2,3866	77,54
	Duri	79,9266	5,054139	84,98074
	Belida	79,9266	-1,418	78,5086
	Banyu	79,9266	1,655864	81,58246
	SC	79,9266	-0,42955	79,49705
November	SLC	79,68939	-1,87657	77,81282
	Attaka	79,68939	-2,4416	77,81282
	Duri	79,68939	5,009413	84,6988
	Belida	79,68939	-1,47401	78,21538
	Banyu	79,68939	1,558697	81,24809
	SC	79,68939	-0,5063	79,18309
December	SLC	79,45219	-2,10863	77,34356
	Attaka	79,45219	-2,4967	76,95549
	Duri	79,45219	4,964687	84,41688
	Belida	79,45219	-1,53001	77,92218
	Banyu	79,45219	1,461529	80,91372
	SC	79,45219	-0,58305	78,86914

DOI: https://doi.org/10.29207/resti.v6i6.4590

3.3 *Performance Evaluation/Validation*

The Root Mean Square Error (RMSE) is a measure of accuracy that shows how well individual pairs of forecast and observation values match up on average. In this measure, the prediction model is said to be the best if the RMSE value is 0 (zero). Table 7 displays the analysis of the MAPE value.

Table 7. MAPE Value Range

Range of MAPE	Meaning
< 10%	Forecasting model ability is very good
10-20%	Forecasting model ability is good
20-50%	Forecasting model ability is fair
>50%	Forecasting model ability is poor

The MSE, RMSE, and MAPE are then used to evaluate the formed equation model's performance, as shown in Tables 8 and 9.

Table 8. MSE, RMSE and MAPE calculation results (Year, Month and Date-Brent Variables)

Type of Crude Oil	MSE	RMSE	MAPE	MAPE Value Analysis
SLC, SC, Duri, Belida, Banyu and SC	386.390 200	19.65681	45%	Forecasting model ability is fair

The result of the MAPE value in Table 8 is above 10% because the source of the data obtained, namely the decree of the Minister of Energy and Mineral Resources regarding the determination of ICP, contains 30% incomplete data. In the preprocessing stage, it is done by calculating the average value each year to complete the incomplete data.

Table 9. Calculation results of MSE, RMSE and MAPE (Variable Year, Month and Alpha)

Type of Crude Oil	MSE	RMSE	MAPE	MAPE Value Analysis
SLC	4.87556	2.20807	9%	Forecasting model ability is very good
Attaka	4.98657	2.23306	45%	Forecasting model ability is fair
Duri	21.04034	4.58698	126%	Forecasting model ability is poor
Belida	5.30362	2.30296	33%	Forecasting model ability is fair
Banyu	3.81721	1.95377	150%	Forecasting model ability is poor
SC	5.98450	2.44633	50%	Forecasting model ability is fair

The MAPE value in Table 9 is for Attaka, Duri, Belida, Banyu and SC crude oil, except for SLC, which has a MAPE value exceeding 100%. This MAPE result is caused by the data source obtained from the Decree of the Minister of Energy and Mineral Resources every month for these five crude oils (Attaka, Duri, Belida, Banyu, and SC) are incomplete; both Date-Brent and Alpha data. So at the preprocessing stage, the preliminary data must be completed by finding the average value every year. This average value is assigned to the five types of oil in the absence of data. This calculation resulted in a higher MAPE result. Meanwhile, for the SLC crude oil type in the Minister of Energy and Mineral Resources Decree, the data for each month has a complete SLC value so that in the MAPE calculations for SLC crude oil, it has a value below 10%, and it can be concluded that the predictive model ability is Very Good.

The Ministry of Energy and Mineral Resources (ESDM) has issued regulations regarding the ICP formula consisting of Dated-Brent plus Alpha for the main crude oil price. Every month the Minister of Energy and Mineral Resources issues a decision letter on the pricing of Indonesian crude oil. Based on monthly data, there are special Dated-Brent and Alpha values for the main crude oil, namely SLC, Attaka, Duri, Belida, Banyu and SC. The parameters used in Multiple Linear Regressions to predict oil prices are Dated-Brent and Alpha values. The Multiple Linear Regressions method predicts Dated-Brent and Alpha values for the next months.

4. Conclusion

This study forecasts Indonesian Crude Oil Prices (ICP) using time series data and Multiple Linear Regressions. The results show that the intercept value, variable coefficient 1 and variable coefficient 2 so that the ICP prediction model is obtained for SLC, Attaka, Belida, Duri, Banyu and SC crude oil types. The model produced, on average, has a good level of accuracy in calculating MAPE where for SLC = 9%, Attaka = 45%, Duri = 126%, Belida = 33%, Banyu = 150% and SC = 50%. Based on the MAPE calculation value, the Linear Regression Equation to predict Indonesian Crude Oil Prices (ICP) shows that the model produced by SLC crude oil is very good. Attaka, Belida and SC crude oil yielded fair yields and Duri and Banyu crude oil yielded poor yields.

Acknowledgment

This research supported by Universitas Islam Riau. Thank you very much for supported by UIR.

Reference

[1] G. Khuziakhmetova, V. Martynov, and K. Heinrich, "DSS for Oil Price Prediction Using Machine Learning," vol. 166, no. Itids, pp. 89–94, 2019.

[2] Y. N. Kunang *et al.*, "Analysis and implementation of the Port Knocking method using Firewall-based Mikrotik RouterOS," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 8, no. 4, pp. 1907–5022, 2019.

DOI: https://doi.org/10.29207/resti.v6i6.4590

[3] Y. Chen, Y. Zou, Y. Zhou, and C. Zhang, "Multi-step-ahead Crude Oil Price Forecasting based on Grey Wave Forecasting Method," *Procedia Comput. Sci.*, vol. 91, pp. 1050–1056, 2016.

[4] H. Rong, A. P. Teixeira, and C. Guedes Soares, "Data mining approach to shipping route characterization and anomaly detection based on AIS data," *Ocean Eng.*, vol. 198, p. 106936, 2020.

[5] A. Veno, L. A. Safitri, and T. Prijanto, "Analisis Daya Saing Ekspor Minyak Mentah Indonesia Dibanding dengan Negara Anggota OPEC," *Triangle 1*, vol. 1, no. 1, pp. 16–29, 2020.

[6] C. EL AMRANI and H. GIBET TANI, "Smarter round robin scheduling algorithm for cloud computing and big data," *J. Data Min. Digit. Humanit.*, 2018.

[7] S. Gao and Y. Lei, "A new approach for crude oil price prediction based on stream learning," *Geosci. Front.*, vol. 8, no. 1, pp. 183–187, 2017.

[8] A. Sepp, "Machine Learning for Volatility Trading (Presentation Slides)," *SSRN Electron. J.*, no. May, 2018.

[9] I. H. Witten, "Data Mining Practical Machine Learning Tools and Techniques (Fourth Edition)," *Morgan Kaufmann*. pp. 417–466, 2017.

[10] M. Hussein and Y. Azhar, "Prediksi Harga Minyak Dunia Dengan Metode Deep Learning," *Fountain Informatics J.*, vol. 6, no. 1, pp. 26–34, 2021.

[11] D. Suryani, A. Yulianti, E. L. Maghfiroh, and J. Alber, "Quality Classification of Palm Oil Products Using Naïve Bayes Method," *Sistemasi*, vol. 11, no. 1, p. 251, 2022.

[12] A. Fitri Boy, "Implementasi Data Mining Dalam Memprediksi Harga Crude Palm Oil (CPO) Pasar Domestik Menggunakan Algoritma Regresi Linier Berganda (Studi Kasus Dinas Perkebunan Provinsi Sumatera Utara)," *J. Sci. Soc. Res.*, vol. 4307, no. 2, pp. 78–85, 2020.

[13] Suyanto, *Machine Learning - Tingkat Dasar dan Lanjut*, Pertama. Bandung: Informatika, 2018.

[14] S. Wang, J. Cao, and P. Yu, "Deep learning for spatio-temporal data mining: A survey," *IEEE Trans. Knowl. Data Eng.*, 2020.

[15] G. N. Ayuni and D. Fitrianah, "Penerapan metode Regresi Linear untuk prediksi penjualan properti pada PT XYZ," *J. Telemat.*, vol. 14, no. 2, pp. 79–86, 2019.

[16] A. Izzah and R. Widyastuti, "Prediksi Harga Saham Menggunakan Improved Multiple Linear Regression untuk Pencegahan Data Outlier," *Kinet. Game Technol. Inf. Syst. Comput. Network, Comput. Electron. Control*, vol. 2, no. 3, pp. 141–150, 2017.

[17] A. K. Marandi and D. A. Khan, "An Impact of Linear Regression Models for Improving the Software Quality with Estimated Cost," *Procedia Comput. Sci.*, vol. 54, no. April 2016, pp. 335–342, 2015.

[18] N. Tomasevic, N. Gvozdenovic, and S. Vranes, "An overview and comparison of supervised data mining techniques for student exam performance prediction," *Comput. Educ.*, vol. 143, p. 103676, 2020.

[19] Y. Supriyanto, "Prediksi Harga Minyak Kelapa Sawit Menggunakan Linear Regression Dan Random Forest," *J. Ilm. Wahana Pendidik.*, vol. 8, no. 7, pp. 178–185, 2022.

[20] N. Aghdaei, G. Kokogiannakis, D. Daly, and T. McCarthy, "Linear regression models for prediction of annual heating and cooling demand in representative Australian residential dwellings," *Energy Procedia*, vol. 121, pp. 79–86, 2017.

DOI: https://doi.org/10.29207/resti.v6i6.4590