VOL. 9. NO. 2 FEBRUARY 2024
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480 /jitk.v9i2.4381

JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

PREDICTION OF BIODIESEL FUEL PRICES USING MULTIPLE LINEAR REGRESSION ALGORITHMS

Deny Haryadi[1*]; Dewi Marini Umi Atmaja[2]; Adi Kuncoro[3]

Information Technology Campus Jakarta[1,3]
Universitas Telkom[1,3]
https://telkomuniversity.ac.id/[1,3]
denyharyadi@telkomuniversity.ac.id[1*], 1202218016@student.telkomuniversity.ac.id[3]

Digital Business [2]
Universitas Medika Suherman[2]
https://medikasuherman.ac.id[2]
dewi@medikasuherman.ac.id[2]

(*) Corresponding Author

Abstract—*Biodiesel is a fuel derived from palm oil and a type of fuel that is an alternative to renewable energy, can be renewed and has the potential to become a substitute for fossil sources that are used non-stop. The utilize of biodiesel can be an arrangement for Indonesia to diminish reliance on imported diesel fuel since biodiesel does not contain sulfur and is demonstrated to be ecologically inviting. The price of biodiesel-type biofuels can increase, decrease, or remain constant due to factors that influence it, including the price of biodiesel competitors, palm oil, and world crude oil. For this reason, it is necessary to have a method that can predict the price of biodiesel-type fuel so that in the future, the price of biodiesel-type biofuel does not decrease or become unable to compete with its competitors. Prediction of biodiesel fuel prices can be done by implementing a multiple linear regression algorithm, one of the data mining algorithms. RMSE results obtained in this study were 0.003 with a standard deviation of +/- 0.000 so it can be concluded that this algorithm is quite accurate in predicting the price of biodiesel-type biofuels. A comparison of the results of manual calculations with the implementation of RapidMiner in the study obtained the same results because there was a causal relationship between attributes. The use of the multiple linear regression algorithm in this research is useful in planning the right strategy and making decisions to maintain biodiesel market price stability in the future.*

Keywords: *biodiesel price, data mining, linear regression algorithm, prediction.*

Intisari—*Biodiesel adalah bahan bakar yang berasal dari minyak kelapa sawit dan salah satu jenis bahan bakar yang merupakan alternatif energi terbarukan, dapat diperbaharui serta berpotensi menjadi pengganti sumber fosil yang terus menerus digunakan. Penggunaan biodiesel merupakan solusi bagi Indonesia dalam mengurangi ketergantungan pada impor solar karena biodiesel tidak mengandung belerang dan terbukti ramah lingkungan. Harga bahan bakar jenis biodiesel dapat naik, turun atau tetap karena faktor-faktor yang mempengaruhinya, antara lain harga kompetitor biodiesel, minyak sawit, dan minyak mentah dunia. Untuk itu diperlukan suatu metode yang dapat memprediksi harga bahan bakar jenis biodiesel agar kedepannya harga bahan bakar nabati jenis biodiesel tidak turun atau kalah bersaing dengan kompetitornya. Prediksi harga bahan bakar biodiesel dapat dilakukan dengan mengimplementasikan algoritma regresi linear berganda yang merupakan salah satu algoritma data mining. Hasil RMSE yang diperoleh pada penelitian ini adalah 0,003 dengan standar deviasi +/- 0,000 sehingga dapat disimpulkan bahwa algoritma ini cukup akurat dalam memprediksi harga bahan bakar nabati jenis biodiesel. Perbandingan hasil perhitungan manual dengan implementasi RapidMiner pada penelitian memperoleh hasil yang sama karena terdapat hubungan sebab akibat antar atribut. Penggunaan algoritma regresi linear berganda dalam penelitian ini bermanfaat dalam merencanakan strategi yang tepat dan pengambilan keputusan untuk menjaga stabilitas harga pasar biodiesel di masa depan.*

Kata Kunci: *harga biodiesel, data mining, algoritma regresi linear, prediksi.*

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 9. NO. 2 FEBRUARY 2024
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v9i2.4381

INTRODUCTION

Renewable energy sources are natural, freely used and renewable energy sources that are continuously, without limitation. One of the newest energy sources is Biodiesel. Biodiesel is one of the most promising energies as a substitute for fossil energy (BBM) derived from the processing of oil-producing plants, therefore biodiesel is often called green energy because its origins and emissions are environmentally friendly and do not significantly increase global warming. Previous research related to Analysis of influencing factors biodiesel crude palm oil offer in Indonesia 2006-2018. The results showed that there was a significant positive relationship between fuel prices, biodiesel prices, and the amount of domestic biodiesel consumption to the total supply of crude palm oil biodiesel in Indonesia [1].

Indonesia is a country that produces palm oil which is often referred to as the largest Crude Palm Oil (CPO) in the world, which is the main raw material for biodiesel [2]. Type of biofuel Biodiesel plays an important role in the national economy, especially government tax revenue [3]. The price of palm oil can rise, fall or stay constant on a daily basis due to factors affecting it, including other vegetable oils (soybean or rapeseed oil), world crude oil, as well as exchange rates. real exchange rate between the dollar and the US dollar. currency. of the country of manufacture (rupiah, ringgit, and Canada) or consumer country currencies (rupees) [4]. Likewise, the price of biodiesel-type biofuels can increase, decrease or remain constant due to factors that influence it, including the price of biodiesel competitors, palm oil, and world crude oil. Therefore, a model is needed that can help predict the price of biodiesel so that relevant agencies can make the right policies in the trading system for the biodiesel-type biofuel industry [5].

The field of data mining science is here as a solution to overcome these problems. Data mining is a process of extracting or extracting previously unknown data, but understandable and useful from large databases and used to make decisions [6]. Linear regression algorithm is one of the data mining algorithms applied to solve prediction problems. This study was conducted to predict biodiesel price by linear regression algorithm, from which the results of this study can be used as a benchmark and a tool to make decisions in forecasting biodiesel price. learn. The attributes used in this study are factors that influence biodiesel prices, such as the prices of biodiesel competitors (diesel prices), palm oil prices, and global crude oil prices.

MATERIALS AND METHODS

A. Renewable Energy

Renewable energy source is a natural energy source, freely used directly and renewable continuously, without limitation, one of the newest energy sources is Bioenergy which is used as electricity, transportation fuel, and heat energy. The newest types of energy are Hydropower, Biomass Power, Geothermal, Solar Energy, Wind Power, Tidal Energy, and Ocean. Thermal Power (Ocean Thermal Energy).

New energy and renewable energy are very important for sustainable development, especially in overcoming the problem of inequality [7]. This can have an impact on its usefulness for development based on energy security, namely, guaranteed energy supply conditions, public access to affordable energy in the long term, and not affected by regional and international turmoil [8].

B. Biodiesel

Biodiesel or fatty acid methyl esters (FAME) are alternative fuels to replace fossil oil (diesel) or bioenergy made from vegetable or animal oils. Vegetable oil as the main source of biodiesel can be obtained from various types of plants that depend on the main resources of a place or country. The development of biodiesel in Indonesia started from the energy crisis that occurred and the decreasing oil reserves in Indonesia [9].

The government has issued the Presidential Decree of the Republic of Indonesia No. 5 of 2006 on National Energy Policy to implement the optimization of fuel supply on a new basis, Renewable Energy (EBT), which among other things increases the use of biofuels (biofuels), namely biodiesel and bioethanol. Previous research related to Association Classification Testing in Improving the Quality of Palm Oil as a Basic Material for Biodiesel. In this research using Data Mining application so that can be used to display the information of quality improvement of palm oil acid with high quality as a biodiesel base where the information is seen from the value of support and Confidence between items. M-Apriori algorithm cant read the continous data, so it must be done transformation first, one of them using the help of discreetization technique. Crude palm oil acid type esteridication which is fatty acid levels > 60 and Abu bunches 51-100 g/mol then the quality of biodiesel with the result of rising methyl Ester increases by as much as 100% (confidence: 1). From data that is calculated manually, no relationship is found Association. [10].

C. Data Mining

Data mining uses artificial intelligence, statistics, and mathematics to extract new knowledge and information from large amounts of data in data warehouses [11]. Data mining may be an innovation that's anticipated to bridge the communication between information and clients [12]. Some of the solutions provided by data mining include:

a) Guess your target market. In data mining, the buyer model can be grouped (clustered) to classify each buyer according to desired characteristics.
b) See buying habits over time. Data mining can be used to see buying patterns over time.
c) Cross-market analysis. Data mining allows you to see the relationship between one product and another.
d) Customer information. Data mining allows users to view buyer profiles, making it more likely that a particular group of buyers will gravitate to a particular product.
e) Summary information. Through data mining, you can create summary reports with multiple dimensions and additional statistical information.

D. Multiple Linear Regression

Multiple linear regression is a regression analysis that describes the relationship between a response variable (dependent variable) and factors that influence multiple predictor variables (independent variables) [13]. When an outcome/output or class is numeric, and all attributes are numeric, linear regression is the fastest technique to solve.

The least squares method is the most common method for setting up simple linear regression equations. The common shape of Simple Linear Regression is as in formula (1). Meanwhile, to determine the values of a and b, the Simple Linear Regression Equation is as in formula (2).

$$b = \frac{n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \quad (1)$$

$$a = \frac{\sum_{i=1}^{n} y_i}{n} - b\frac{\sum_{i=1}^{n} x_i}{n} \quad (2)$$

N : the number of data pairs
Yi : value of the i-th dependent variable Y
Xi : value of the i-th independent variable X

Calculate the linear regression equation.

$$y = a + bx \quad (3)$$

Y : dependent variable
A : Constant
X : independent variable
B : Slope

Multiple regression analysis allows researchers to predict how the dependent variable (reference) will change (increase or decrease) when two or more independent variables are manipulated as predictors (decrease or increase in value) [14]. Therefore, multiple regression analysis is performed when there are two or more independent variables. The regression equation for the two predictors is as in formula (4).

$$y = a + b1x1 + b2x2 \quad (4)$$

The regression equation for the three predictors is as in formula (5)

$$y = a + b1x1 + b2x2 + b3x3 \quad (5)$$

The regression equation for n predictors is as in formula (6)

$$y = a + b1x1 + b2x2 + b3x3 + b4x4 \quad (6)$$

To be able to make predictions through regression, the data for each variable must be available. Furthermore, based on the data the researcher must be able to find the regression equation through calculations.

E. Root Mean Squared Error

Root Mean Squared Error or RMSE is a squared scoring rule that also measures the size of the average error. RMSE is the root mean square of the difference between predictions and actual observations [15]. The way that is also usually used to evaluate data mining models is the RMSE. This method is also known as the root mean squared deviation (RMSD). As can be guessed from its name, RMSE or RMSD is calculated by squaring the error (predicted – observed) divided by the amount of data (= average), then taking its roots. Mathematically the formula is written as in formula (7).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(yi^1 - yi)^2} \quad (7)$$

F. Thinking Framework

The framework of thought is useful as a guide so that research can be carried out consistently. In this study is used multiple linear regression

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 9. NO. 2 FEBRUARY 2024
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v9i2.4381

algorithm. To solve this problem, performance tests are performed on the model generated from the algorithm. The thinking framework can be seen in Figure 1.

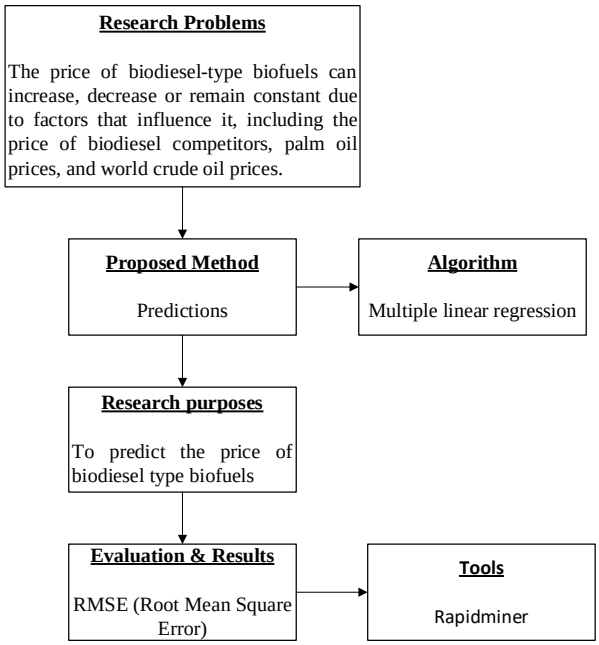

Figure 1. Thinking Framework

G. Data Collection

The information utilized in this think about was taken from the official site www.investing.com and contains an add up to 1000 data. The survey data consists of the attributes or variables that influence them, such as competitive prices for biodiesel (diesel prices), palm oil prices, and global crude oil prices from 2012 to 2022. Variable definitions for biodiesel data can be seen in Table 1.

Table 1. Defening Variables for Biodiesel Data

Variable	Definition
Palm oil	Price of palm oil
Crude oil	World crude oil prices
Solar	Biodiesel competitor prices
Biodiesel	Biodiesel prices

RESULTS AND DISCUSSION

A. Preprocessing Data

Data preprocessing transforms raw data into a dataset by cleaning, selecting, and transforming the data so that it can be more easily accepted by a variety of linear regression algorithms. In this study, the data preprocessing phase includes determining the inputs and outputs of dependent and independent variable, normalizing the data, and sharing the testing and training data.

a) Determination of dependent and independent variables. The data used for this study are palm oil price data, world crude oil price data, biodiesel competitor price data (diesel prices), and biodiesel price data from 2012–2022 including 1000 data. This research applies a type of learning called supervised learning. Guided learning requires learning input and output variables using an algorithm. The input data used in this study are the factors affecting the price of biodiesel type biofuel as shown in Table 2.

Table 2. Input and Output Patterns of Biodiesel Data

No	X_1 (Palm oil)	X_2 (Crude oil)	X_3 (Solar)	Y (Biodiesel)
1	2,428	49,16	125,08	0,0009
2	2,435	49,59	127,84	0,0009
3	2,444	49,03	127,55	0,0009
...	...	...	...	...
...	...	...	...	...
...	...	...	...	...
998	6,393	65,37	49,55	0,0017
999	6,546	66,27	47,75	0,0017
1000	6,489	65,49	48,21	0,0013

b) Normalize the data. In this phase, input and output data is normalized between zero and one using min-max normalization tool.

c) Train data and test data. We split the data into testing and training data and perform analysis using multiple regression algorithms. Shared training data is used to improve the performance of linear regression algorithms and test data to determine optimal parameters for model building. The data distribution can be seen in Table 3.

Table 3. Share Training Data and Testing Data

Information	Data		Total
	Training	Testing	
Amount of Data	900	100	1000
Percentage	90%	10%	100%

From table 3 we can see that the data exchange used in this study accounted for 90% of the total data used as practice/exercise data, the remaining 10% as test/test data experience. As machine learning trains better to learn the model, the training data volume percentage is larger. This is done so that machine learning can inform the model, and the former model can be trained using data tests to provide data test predictions. The training/exercise data and test/trial data distributions were randomly distributed using Rapid Miner. In addition, the training data will be trained using multiple linear regression algorithms so that the model is trained with the combination

of parameters used and then tested the data including the test results. The output of the model is trained from the training data.

B. Linear Regression Analysis

Modeling in this study was carried out by data mining prediction techniques using a linear regression algorithm. Linear regression is used to explore how the dependent variable can be independently predicted by the independent variable or the predictor variable. The regression analysis effect can be used to increase or decrease the dependent variable by increasing or decreasing the state of the independent variable or by increasing or decreasing the state of the dependent variable relative to the independent variable or vice versa. Multiple regression analysis allows researchers to predict how the dependent variable (reference) will change (increase or decrease) when two or more independent variables are manipulated as predictors (increase or decrease in value). used if you want to. Therefore, multiple regression analysis is performed when the number of independent variables increases.

$$b_1 = \frac{n\sum(x_1y)-(\sum x_1)(\sum y)}{n(\sum {x_1}^2)-(\sum x_1)^2} \quad (8)$$

$$b_1 = \frac{100(0)-(234)(0)}{100(549)-(234)^2}$$

$$b_1 = \frac{0-0}{54900-54756}$$

$$b_1 = \frac{0}{144}$$

$$b_1 = 0$$

$$b_2 = \frac{n\sum(x_2y)-(\sum x_2)(\sum y)}{n(\sum {x_2}^2)-(\sum x_2)^2} \quad (9)$$

$$b_2 = \frac{100(0)-(5211)(0)}{100(272929)-(5211)^2}$$

$$b_2 = \frac{0-0}{27292900-27154521}$$

$$b_2 = \frac{0}{138379}$$

$$b_2 = 0$$

$$b_3 = \frac{n\sum(x_3y)-(\sum x_3)(\sum y)}{n(\sum {x_3}^2)-(\sum x_3)^2} \quad (10)$$

$$b_3 = \frac{100(0)-(8168)(0)}{100(707597)-(8168)^2}$$

$$b_3 = \frac{0-0}{70759700-66716224}$$

$$b_3 = \frac{0}{4043476}$$

$$b_3 = 0$$

$$a = \frac{\sum y - b_1\sum x_1 + b_2\sum x_2 + b_3\sum x_3}{n} \quad (11)$$

$$a = \frac{0-((0)*(234))+((0)*(5211))+((0)*(8168))}{100}$$

$$a = \frac{0-0+0+0}{100}$$

$$a = \frac{0-0}{100}$$

$$a = \frac{0}{100}$$

$$a = 0$$

Testing data can be viewed in Table 4.

Table 4. Biodiesel Price Prediction

No	X1 (Palm Oil)	X2 (Crude Oil)	X3 (Solar)	Y (Biodiesel)
1	2,188	57,16	78,86	?
2	2,197	57,46	77,61	?
3	2,180	58,09	78,54	?

Calculating Linear Regression Equations

$$y = 0 + (0 \text{ x } 2{,}188) + (0 \text{ x } 57{,}16) + (0 \text{ x } 78{,}86) = 0$$

$$y = 0 + (0 \text{ x } 2{,}197) + (0 \text{ x } 57{,}46) + (0 \text{ x } 77{,}61) = 0$$

$$y = 0 + (0 \text{ x } 2{,}180) + (0 \text{ x } 58{,}09) + (0 \text{ x } 78{,}54) = 0$$

C. Evaluation of Linear Regression Model

The next step after the linear regression analysis process is to perform data modeling in Rapid Miner. The models used in Rapid Miner use a linear regression algorithm as shown in Figure 2.

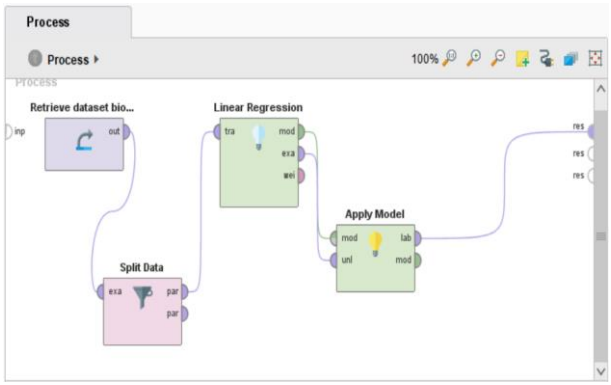

Figure 2. Model Evaluation Process on RapidMiner

The next step is to import the training data and test the data to be tested to make predictions

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 9. NO. 2 FEBRUARY 2024
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v9i2.4381

about the properties of the class can be seen in Figure 3.

Row No.	biodiesel (y)	prediction(bi...	minyak kela...	minyak men...	solar (x3)
1	0.001	0.005	2.428	49.160	125.080
2	0.001	0.005	2.435	49.590	127.840
3	0.001	0.005	2.444	49.030	127.550
4	0.001	0.005	2.458	49.580	122.310
5	0.001	0.005	2.453	49.390	121.690
6	0.001	0.005	2.432	49.170	121.750
7	0.001	0.005	2.437	49.560	121.690
8	0.001	0.005	2.435	48.590	118.440
9	0.001	0.005	2.423	48.820	114.420
10	0.001	0.005	2.410	47.590	114.480
11	0.001	0.005	2.429	47.550	115.410
12	0.001	0.005	2.461	46.780	117.820
13	0.001	0.005	2.437	47.090	116.300
14	0.001	0.005	2.455	48.510	116.990
15	0.001	0.005	2.469	47.370	118.220

Figure 3. Prediction Results

After examining the predictions, the next step is to measure the accuracy of the prediction results. This can be seen in Figure 4.

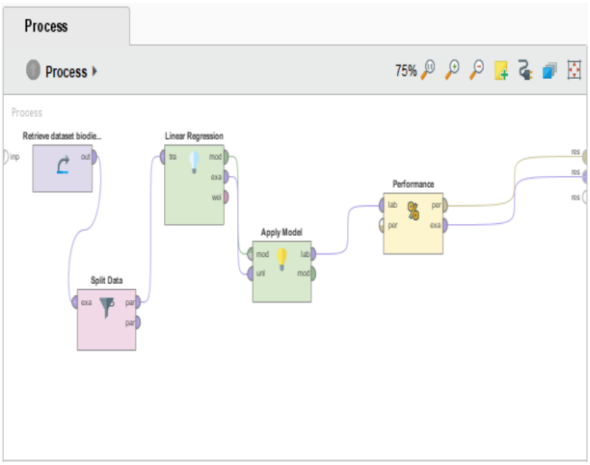

Figure 4. Search Process Root Mean Squared Error

To form running information less demanding to studied, it is vital to purport execution devices to discover the root mean squared error. The result can be seen in Figure 5.

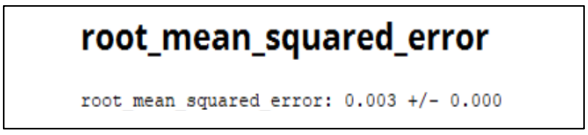

Figure 5. Root Mean Square Error Test Results

The moment step is to actualize the direct relapse calculation utilizing rapid-miner apparatuses. The taking after are the stages in applying the straight relapse calculation:

a) Determine predictions on test data run by Rapid Miner and create prediction confidence values.
b) Determine the output power and find the root mean square error.

A split-validation demonstrate comprises of two parts: a preparing segment (utilized for the classification calculation) and a testing area (containing a show application work that applies the demonstrate to information tests and a execution work that shows the root cruel square blunder).

D. Biodiesel Price Prediction

After calculating the whole values X1Y, X2Y, X3Y, X1X2X3, X12, X22, X32 of variable X1 in 100 records of the preparing information set, the full result of each esteem is 234. The variable is 5211, the entirety of the factors is X3 8168, the entirety of the factors Y is 0, the entirety of X1Y is 0, the whole of X2Y is 0, the entirety of the variables is 0, the whole of X1X2X3 is 550, the whole of X12 is The whole esteem is 549, the X22 add up to esteem is 272929, and the X32 add up to esteem is 707597. Calculate the values of factor b1, factor b2, factor b3, and factor a from the sum above and transform the formula so that the value of b1 is 0, the value of b2 is 0, and the value of b3 is 0, and the value of a is 0. It can be obtained. These three coefficient values are used in the multi-equation application of the linear regression algorithm. From the values of the above coefficients, we can use a simple mathematical model to find the equations for the linear regression algorithm of the model Y = 0 + (0.X1) + (0.X2) + (0.X3). Variable X1 corresponds to data on coconut and palm oil prices, variable X2 corresponds to world crude oil prices, and variable X3 corresponds to diesel oil prices.

A twofold direct relapse will at that point be performed to foresee the exploratory information. By and large, the application of the straight relapse condition is connected to foresee the 3 information set records that are predefined test information. From the calculations performed physically additionally concurring to the method within the RapidMiner application, the shown comes about don't contrast altogether, in other words, the manual calculation and the calculation processed within the application for comparative comes about. A comparison of the results of the manual calculations and the results of the RapidMiner application for the variable Y is shown in Table 5.

Table 5. Manual Comparison and Fast Mining

No	X1 (Palm Oil)	X2 (Crude Oil)	X3 (Solar)	Y (Biodiesel) Manual	Y (Biodiesel) Rapidminer
1	2,188	57,16	78,86	0	0
2	2,197	57,46	77,61	0	0
3	2,180	58,09	78,54	0	0

Comparing the genuine esteem of variable Y from test information (perception) with the esteem of variable Y anticipated in RapidMiner application, can be seen in Figure 6.

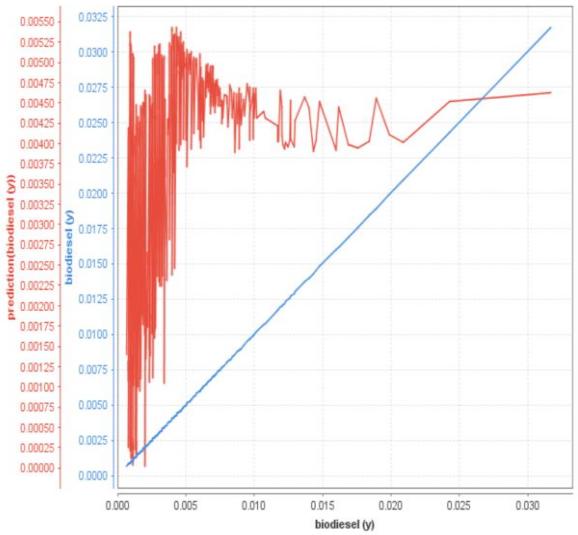

Figure 6. Graph of Comparison of (Y) Observation with (Y) Predictions

In general, you can see from the graph that the actual values of the variable Y are indicated by the blue lines, and the predicted values of the variable Y are indicated by the red lines. From this data, the predicted values are in the same range as the actual Y values (observed). This histogram is evaluated using the Root Mean Square Error method to determine the size of the error values. Another evaluation of this model is to find the original mean squared error or initial Root Mean Square Error value. Using the RapidMiner app, the Root Mean Square Error value is 0.003 with a standard deviation of +/- 0.000.

The variables used in this study are determined based on the results of tests performed (palm oil price, world crude oil price, diesel oil price, biodiesel price) are demonstrated by use a linear regression algorithm which has a big impact on good performance in terms of mean squared error and you can get 0.003 +/- 0.000. In fact, there is a correlation relationship (causality) between one variable (independent or predictor) and another variable (dependent or criterion). This experimental process is performed to predict the price of biodiesel-type biofuels using a multiple linear regression algorithm.

CONCLUSION

Based on the results of this study, the following conclusions can be drawn: In this study, attributes that affect biodiesel prices are used, such as the prices of global biodiesel, palm oil and crude oil competitors. From these attributes, a linear regression algorithm is applied to predict the biodiesel price based on the functional relationship with the data attributes. The process of predicting biodiesel price data using a linear regression algorithm starts from data processing, linear regression analysis, evaluation of linear regression models, and biodiesel price predictions. The use of a linear regression algorithm is demonstrated based on the results of the tests performed, that the attributes used in this study have a significant influence on this study. This can be done using the root mean square with a value of 0.003 +/- 0.000 giving good results. In addition, comparing the results of the manual calculation with the RapidMiner implementation in the study yielded similar results because there is a causal relationship between the attributes. This can be seen in the comparison of biodiesel price predictions using manual calculations with the RapidMiner implementation where from the price of palm oil = 2,188, the world crude oil price = 57.16, and the price of solar = 78.86, we get the same biodiesel price prediction, namely 0.0019 in August 2022.

REFERENCE

[1] E. Wijiastuti, H. Sasana, and G. Jalunggono, "Analisis Faktor-Faktor Yang Mempengaruhi Penawaran Biodiesel Crude Palm Oil Di Indonesia Tahun 2006-2018," *Din. Dir. J. Econ.*, vol. 2, no. 3, pp. 616–631, 2020, doi: 10.31002/dinamic.v2i3.1414.

[2] V. Nurmalita and P. Ari Wibowo, "Analisis Faktor-Faktor yang Mempengaruhi Ekspor Minyak Kelapa Sawit Indonesia ke India.," *Econ. Educ. Anal. J.*, vol. 8, no. 2, pp. 605–619, 2019, doi: 10.15294/eeaj.v8i2.31492.

[3] P. Vila Tri Wahyuni, S. Wahyuni Mustafa, and R. Solling Hamid, "Pengaruh Harga Internasional dan Nilai Tukar terhadap Permintaan Ekspor Minyak Sawit di Indonesia," *Jesya (Jurnal Ekon. Ekon. Syariah)*, vol. 4, no. 2, pp. 1104–1116, 2021, doi: 10.36778/jesya.v4i2.420.

[4] A. Rahman Hakim, D. Marini Umi Atmaja, A. Basri, and M. Syafii, "Implementation of Random Forest Algorithm on Palm Oil Price Data," *J. Tech-E*, vol. 6, no. 2, pp. 34–42, 2023, doi: https://doi.org/10.31253/te.v6i2.1757.

[5] I. Solihat, J. Jasmi, and F. Fachruddin,

VOL. 9. NO. 2 FEBRUARY 2024
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v9i2.4381

"Prediksi Jumlah Produksi Kelapa Sawit Dengan Regresi," *SNASTIKOM (Seminar Nas. Teknol. Inf. Komunikasi)*, vol. 1, no. 1, pp. 17–26, 2020.

[6] E. Triyanto, H. Sismoro, and A. Dwi Laksito, "Implementasi Algoritma Regresi Linear Berganda Untuk Memprediksi Produksi Padi Di Kabupaten Bantul," *Rabit J. Teknol. dan Sist. Inf. Univrab*, vol. 4, no. 2, pp. 73–86, 2019, doi: 10.36341/rabit.v4i2.666.

[7] W. Strielkowski, E. Tarkhanova, M. Tvaronavičˇ, and Y. Petrenko, "Renewable Energy in the Sustainable Development of Electrical," *Energies*, vol. 14, pp. 1–24, 2021, doi: https://doi.org/10.3390/en14248240.

[8] R. Advent, Z. Zulgani, and N. Nurhayani, "Analisis faktor - faktor yang mempengaruhi ekspor minyak kelapa sawit di Indonesia Tahun 2000-2019," *e-Journal Perdagang. Ind. dan Monet.*, vol. 9, no. 1, pp. 49–58, 2021, doi: 10.22437/pim.v9i1.13652.

[9] A. Pradana, K. Edy Sularso, and I. Kartika Eka Wijayanti, "Analisis Faktor yang Berpengaruh terhadap Ekspor Biodiesel di Uni Eropa," *J. READ (Research Empower. Dev.*, vol. 1, no. 1, pp. 41–50, 2021, doi: https://doi.org/10.20884/1.read.2021.2.1.3531.

[10] A. Fitra Andikos and H. Andri, "Pengujian Association Clasification Dalam Meningkatkan Kualitas Minyak Sawit Sebagai Bahan Dasar Biodisel," *J. Media Inform. Budidarma*, vol. 3, no. 4, pp. 340–345, 2019, doi: 10.30865/mib.v3i4.1383.

[11] 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. 3, no. 2, pp. 78–85, 2020, doi: https://doi.org/10.54314/jssr.v3i2.421.

[12] D. Haryadi and R. Mandala, "Prediksi Harga Minyak Kelapa Sawit Dalam Investasi Dengan Membandingkan Algoritma Naïve Bayes, Support Vector Machine dan K-Nearest Neighbor," *IT Soc.*, vol. 4, no. 1, pp. 28–38, 2019, doi: 10.33021/itfs.v4i1.1181.

[13] A. Al-Fadhilah Nur Wahyudin, A. Primajaya, and A. Susilo Yuda Irawan, "Penerapan Algoritma Regresi Linear Berganda Pada Estimasi Penjualan Mobil Astra Isuzu," *Techno.Com*, vol. 19, no. 4, pp. 364–374, 2020, doi: 10.33633/tc.v19i4.3834.

[14] D. Sari Oktavia Panggabean, E. Buulolo, and N. Silalahi, "Penerapan Data Mining Untuk Memprediksi Pemesanan Bibit Pohon Dengan Regresi Linear Berganda," *JURIKOM (Jurnal Ris. Komputer)*, vol. 7, no. 1, pp. 56–62, 2020, doi: 10.30865/jurikom.v7i1.1947.

[15] D. Haryadi and D. Marini Umi Atmaja, "Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Tingkat Risiko Penyakit Jantung," *J. Informatics Commun. Technol.*, vol. 3, no. 2, pp. 51–66, 2021, doi: 10.52661/j_ict.v3i2.85.