JOURNAL OF DATA SCIENCE | Vol.2023:24
eISSN:2805-5160

Multiple Linear Regression for Predicting the Ship Booking Time: A Case Study at PT. Samudera Indonesia

Tri Basuki Kurniawan[1], Aldi Alvino[2]

[1,2]Faculty of Computer Science, Bina Darma University

[*]**Email:** [1]tribasukikurniawan@binadarma.ac.id, aldiialvino@gmail.com

Abstract

A statistical method called multiple linear regression (MLR), or just multiple regression, makes use of many explanatory variables to forecast the value of a response variable. We studied PT. Samudera Indonesia, a company in the shipping sector, for this paper. One of the companies providing services for maritime transportation is this one, which deals with the inflow and outflow of commodities. Our study focuses on the application of ship docking time prediction at PT. Samudera Indonesia, which is situated at Boom Baru port in Palembang. This research makes use of historical data on the ship's docking time during the preceding three (3) years, utilizing multiple linear regression techniques. Using the company's dataset for the years 2018 to 2020 which consists of 70% training data and 30% testing data the experiment was conducted and recorded. The model's performance has yielded very positive results, as evidenced by the 1.132 RSME (root mean square error) number, 1.075 absolute error, and 1.01% relative error. These numbers closely matched the initial figures that the company had documented.

Keywords

Data Mining, Production Prediction, Multiple Linear Regression Algorithm

Introduction

A port is a wave-protected body of water where ships and other water vehicles can berth, raise, and lower people and cargo, repair, refuel and perform other operations. It has a pier where ships can tie up. Taps are utilized in transit warehouses, locations where items are kept longer before being distributed to the areas, and for the loading and unloading of cargo (Triadmodjo, 2009).

By facilitating the movement of commodities in and out of the industrial sector, PT. Samudera Indonesia is a loading and unloading shipping service provider. The ability and speed with which items are handled during the unloading process from the ship to the warehouse or stacking yard, or the loading process from the warehouse or stacking yard to the ship, is known as the loading and unloading productivity (Saut Gurning & Budiyanto, 2007). Several indications show the level of that ability, including 1) the average number of loading and unloading operations performed per hour by one (1) gang of workers (12 people) on board the ship, measured in

Submission: 27 November 2023; **Acceptance:** 4 December 2023

tons/gang/hour (T/G/J); and 2) the average number of loading and unloading operations performed per hour by one (1) gang of workers (12 people) on board the ship. 2) the average amount of loading and unloading of goods accomplished per hour and carried out by container loading and unloading equipment (B/C/H); 3) the average amount of loading and unloading of goods accomplished per hour and carried out by all third gangs on board while the ship is at the dock (BWT), which is measured in tons/ship/hour (T/K/J) and is commonly referred to as "ship's output".

Data mining is a technique of digging up hidden or valuable information in an extensive data collection (database) so that an interesting pattern that was previously unknown is found. Get correct, new, practical, and understand a style or design from the data. Data mining is a series of processes to manually extract added value in the form of information that has so far been unknown from a database. The information obtained is obtained by extracting and recognizing important or interesting patterns from the data contained in the database (Kusrini & Luthfi, 2009), which is consistent with the research conducted by Muzakir & Wulandari (2016) titled Data Mining Model as Prediction of Hypertension in Pregnancy with the Decision Tree Technique, which states that the implementation of data mining with the decision tree technique using the C4.5 algorithm can produce information.

The Prediction function is a minor data mining function. Prediction is a method of predicting what will happen in the future. Forecasting methods are used in data mining technologies to predict future possibilities based on evidence obtained in data. Prediction is linked to mathematical and statistical calculations (Laim & Kumar, 2023). The mathematical linear regression formula can be used to do prediction stages. As a phase in the predictive analytic process, there are two (2) types of regression formulas: basic linear regression and multiple linear regression (Hariaji, 2021). Linear regression measures predictive data through a straight line to illustrate the correlation relationship between 2 or more variables. Linear regression prediction is used to study how the variables relate to the data forecasting process. How to read the relationship correlation hypothesis data results of prediction analysis is done through statistical analysis.

Multiple linear regression is a regression analysis that explains the relationship between the dependent variable and factors that affect more than one independent variable. Simple and multiple linear regression are almost the same, but multiple regression of the independent variable has more than one predictor variable. Multiple linear regression analysis aims to measure the intensity of the relationship between two or more variables. Multiple linear regression is a regression algorithm that involves more than one independent variable or multiple regression model (Luh et al., 2015). It has the goal of predicting or estimating the value of variable Y. Besides that, it also considers the importance of other variables that can affect variable Y. Therefore, there is one dependent variable, Y, with different independent variables, X1, X2, and Xn. To predict Y, if the values of all independent variables are known, then the multiple linear regression equation can be applied, where the relationship Y and X*1*, X*2*, . . ., X*n* is as follows:

$$(Y = a + b1X1 + b2X2 + \ldots + bnXn) \quad (1)$$

Information:

Y = Dependent variable (predicted value) X1, X2 = Independent variables	a = Constant (Y value if X*1*, X*2*, …, X*n* = 0) b = Regression coefficient (increasing or decreasing value)

Methodology

The data to be analyzed is on ship visits by flag from 2018 to 2021 at PT. Samudera Indonesia goes through the KDD (Knowledge Discovery in Database) stage (Akanmu & Jaja, 2012). The explanation of several steps in the methodology that we carried out is outlined below.

Step 1. Data Selection

The data used in this study comes from data on ship visits by flag from 2018 to 2020 at PT. Samudera Indonesia, this data the researchers obtained through staff employees of PT. Samudera Indonesia and data in the form (Excel) received before the selection process can be seen in Figure 1 below.

NAMA PERUSAHAAN : PT.SAMUDERA AGENCIES INDONESIA Tbk CABANG PALEMBANG
BIDANG USAHA : PERUSAHAAN PELAYARAN
BULAN : AGUSTUS 2020

NO	NAMA KAPAL / BENDERA	DWT / GRT / LOA	PRINCIPAL	OPS KAPAL / STATUS	BONGKAR: TIBA TGL	BERANGKAT TGL	IMPORT: JENIS CARGO	IMPORT: JUMLAH TON	IMPORT: PELABUHAN ASAL	INTERINSULAR: JENIS CARGO	INTERINSULAR: JUMLAH TON/M3	INTERINSULAR: PELABUHAN ASAL	TEMPAT KEGIATAN	MUAT: EXPORT: JENIS CARGO	EXPORT: JUMLAH TON/M3	EXPORT: PELABUHAN TUJUAN	INTERIN: JENIS CARGO	INTERIN: JUM TON
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	MV.PAC CERGAS	9500	SSL	AG	03,08	04,08	Kontainer	4,670.00	SPore	-	-	-	B. Baru	General Cargo		S.pore		
	210	7542			dalam container :FD20/FD40=288 BX										5,682.00			
		118												Container :FD20/40: 264 Box				
2	MV.PAC CERGAS	9500	SSL	AG	10.08	11,08	Kontainer	5,013.00	SPore	-	-	-	B. Baru	General Cargo		S.pore		
	211	7542			dalam container :FD20/FD40: 211 BX										6,397.00			
		118												Container :FD20/40: 284 Box				
3	MV.PAC CERGAS	9500	SSL	AG	17,08	18,08	Kontainer	5,077.00	SPore	-	-	-	B. Baru	General Cargo		S.pore		
	212	7542			dalam container :FD20/FD40: 215 BX										6,350.00			
		118												Container :FD20/40: 293 Box				
4	MV.PAC CERGAS	9500	SSL	AG	24,08	25,08	Kontainer	4,407.00	SPore	-	-	-	B. Baru	General Cargo		S.pore		
	213	7542			dalam container :FD20/FD40: 260 BX										6,315.00			
		118												Container :FD20/40: 301 Box				
5	MV.PAC CERGAS	9500	SSL	AG	31,08	01,09	Kontainer	4,733.00	SPore	-	-	-	B. Baru	General Cargo		S.pore		
	214	7542			dalam container :FD20/FD40: 283 BX										6,381.00			
		118												Container :FD20/40: 280 Box				
6	MV.INTAN DAYA	4652	SSL	AG	31,08	01,09			SPore	GENCAR	639	JKT	B. Baru	General Cargo		S.pore		
	0178	2998									70 BOX				3,259.00			
		89												Container :FD20/40: 157 Box				

Figure 1. Data before selection

Then, the data selection process is carried out. The researcher will carry out the data selection process by removing unnecessary attributes. The data used in conducting this research is data on ship berthing time from 2018 to 2020 in the form of a monthly presentation. This data has several rows, including the date, month, year, volume flag, and lean time. The above data is raw data originating from PT. Ocean Indonesia directly. The data selection process is done manually by determining the attributes/columns for use in the next stage. Figure 2 shows the data that has been selected.

Bulan	Tahun	Bendera	Volume	Waktu Sandar
Januari	2018	6	9888	145
Februari	2018	6	2357	100
Maret	2018	11	5423	183
April	2018	8	2344	67
Mei	2018	7	8764	35
Juni	2018	6	2234	18
Juli	2018	7	52	23
Agustus	2018	5	55	123
September	2018	7	4344	109
Oktober	2018	13	3335	134
November	2018	11	2393	114
Desember	2018	12	7342	189
Januari	2019	10	2657	35
Februari	2019	9	1208	113
Maret	2019	9	7664	89
April	2019	11	5432	56
Mei	2019	9	8897	60
Juni	2019	12	6445	80
Juli	2019	16	6588	73
Agustus	2019	11	6345	95
September	2019	13	4655	56
Oktober	2019	14	6753	34
November	2019	13	2346	46
Desember	2019	14	9866	18

Figure 2. Data after selection

Step 2. Preprocessing

At this preprocessing stage, a data integration process will be carried out to connect some data. Then, a data cleaning process will be carried out to produce a clean dataset for the next stage, namely mining, to obtain patterns regarding ship visits to predict the number of subsequent visits (Oliveira et al., 2017). The following is an explanation of the two processes above:

Step 2.a. Data Integration

The data integration stage involves merging data from several sources (Lenzerini, 2002). At this stage, data on ship visits by flag from 2018 to 2020 was combined at PT. Indonesian Ocean.

Step 2.b. Data Cleaning

The data cleaning stage is the initial stage of the KDD (Knowledge Discovery In Database) process (Rahman et al., 2019). At this stage, irrelevant, missing values and redundant data must be cleaned.

a) Relevant: This is data that focuses on the conditions to be mined. In this case, the attributes to be drilled must be appropriate.
b) Missing Value: This is empty data, so that we will discard open or unfilled data in the database.
c) Redundant: Represents data that is repeated. Occupation data entered twice or more will be erased; only one will be entered. It is because data that is meaningful has no missing values and is not redundant, which is a prerequisite for conducting data mining. Data is said to be missing value if there are characteristics in the dataset that do not have values or are empty. In contrast, data is said to be redundant if there are several records with the same value in the same dataset.

The results of the cleaning process or data cleaning, the visit data obtained previously amounted to 357 records to 248 data records.

Step 2. c. Transformations

The data transformation stage is changing data into a form suitable for mining. An example of data results ready for data mining is as many as 1120 records to be processed. After all the data has been grouped, the data will be added up to get the results of ship visit data for data mining activities.

Step 2. d. Data Mining

After transforming the data into a data form suitable for the application of data mining, the researcher will then simulate predictive calculations using the Multiple Linear Regression method. An example of calculating the multiple linear regression method for ship visits is shown in Table 1 below.

Table 1. Calculation of the Multiple Linear Regression Method

n (Period)	X1 Flag	X2 Volume	Y Back Time	X1^2	X2^2	Y^2	X1.Y	X2.Y
1	6	3432	88	36	11778624	7744	528	302016
2	6	4423	234	36	19562929	54756	1404	1034982
3	8	4322	342	64	18679684	116964	2736	1478124
4	8	6544	102	64	42823936	10404	816	667488
5	11	7898	98	121	62378404	9604	1078	774004
6	10	6445	56	100	41538025	3136	560	360920
7	11	2346	198	121	5503716	39204	2178	464508
8	13	9086	165	169	82555396	27225	2145	1499190
9	13	4344	214	169	18870336	45796	2782	929616
10	14	7764	180	196	60279696	32400	2520	1397520
11	15	2233	198	225	4986289	39204	2970	442134
12	16	7445	93	256	55428025	8649	1488	692385
Total	**131**	**66282**	**1968**	**1557**	**424385060**	**395086**	**21205**	**10042887**

Resolution:

$\sum$ $= \sum - \frac{(\Sigma\ )}{\quad}$
$= 1557 - (/12) = 126,91$

$\sum$ $= \sum - \frac{(\Sigma\ )}{\quad}$
$= 424385060 - (/12) = 58276433$

$\sum$ $= \sum \frac{(\Sigma\ )}{\quad}$
$= 395086 - (/12) = 72334$

$\sum$ $= \sum \frac{((\Sigma\ )(\Sigma\ ))}{\quad}$
$= 21205 - ((131,1968)/12) = -279$

$\sum$ $= \sum \frac{((\Sigma\ )(\Sigma\ ))}{\quad}$
$= 10042887 - ((66282,1968)/12) = -827361$

$\sum$ $= \sum \frac{((\Sigma\ )(\Sigma\ ))}{\quad}$
$= 8682942 - ((131,66282)/12) = 7959363,5$

b1 $= \frac{((\Sigma\ )(\Sigma\ )(\Sigma\ )(\Sigma\ ))}{((\Sigma\ )(\Sigma\ )(\Sigma\ ))}$
$= (58276433)(-279) - ((7959363,5)(-827361))/$
$((126,1)(58276433) - (7959363,5^2))$
$= (-245,1-60849,583) / (700,7-3706231942,1689)$
$= -32325,229$

b2 $= \frac{((\Sigma\ )(\Sigma\ )(\Sigma\ )(\Sigma\ ))}{((\Sigma\ )(\Sigma\ )(\Sigma\ ))}$
$= (126,91)(-827361) - (7959363,5)(-279)$
$/ (126,91)(58276433) - (7959363,5^2)$
$= (-105000384,51-(2220662416,5)) /$
$(7395862112,03- 63351467325132,25)$
$= -3,339$

a $= \frac{\Sigma - \Sigma - \Sigma}{\quad}$
$= (10042887/12) - (32325.229)(131/12) -$
$(3,339)(66282/12)$
$= 836907,25 - (352862,199764)/ -18442,9665$
$= -64,510$

After doing the calculations, the values for the coefficients are obtained, including:

b1 = -32325.229
b2 = -3.339
a =-64,510

So, the multiple regression equation is Y = -64.510 - (32325.229) - (3.339)

Results and Discussion

We use training data from 2018, 2019, and 2020, which is made up of 70% training and 30% testing data. The property that serves as a label is Berthing Time, which is the source of training and testing data in Excel format for predicting ship berthing time so that the RapidMiner software may access it. The data should next be imported into RapidMiner, namely the training and testing data. As illustrated below in Figure 3:

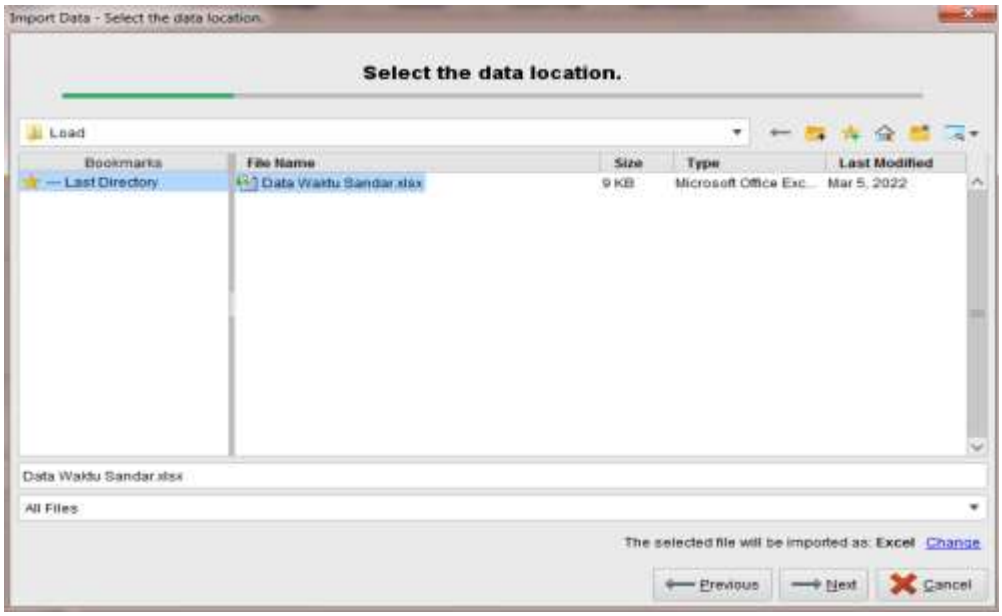

Figure 3. Data Import to RapidMiner

Results

The data to be processed will be divided into two parts: training data and testing data.

Data Training

This training data is 70% of the total data in the View RapidMiner window, and the training data will appear. Like Figure 4 below

Open in Turbo Prep Auto Model Filter (36 / 36 examples): all

Row No.	Waktu Sandar	Bulan	Bendera	Volume
1	145	Januari	6	9888
2	100	Februari	6	2357
3	183	Maret	11	5423
4	67	April	8	2344
5	35	Mei	7	8764
6	18	Juni	6	2234
7	23	Juli	7	52
8	123	Agustus	5	55
9	109	September	7	4344
10	134	Oktober	13	3335
11	114	November	11	2393
12	189	Desember	12	7342
13	35	Januari	10	2657
14	113	Februari	9	1208

Figure 4. Data Training 70%

Data Testing

This testing data is 30% of the total data in the View RapidMiner window training data will appear, as shown in Figure 5.

Open in Turbo Prep Auto Model Filter (12 / 12 examples): all

Row No.	Waktu Sandar	Bulan	Bendera	Volume
1	88	Januari	16	3432
2	234	Februari	13	4423
3	342	Maret	12	4322
4	102	April	13	6544
5	98	Mei	13	7898
6	56	Juni	11	6445
7	198	Juli	13	2346
8	165	Agustus	13	9086
9	214	September	19	4344
10	180	Oktober	17	7764
11	198	November	20	2233
12	93	Desember	19	7445

Figure 5. Data Testing 30%

Multiple Linear Regression Models

Formation of Predictive Models

Figure 6 shows the prediction model of the linear regression method.

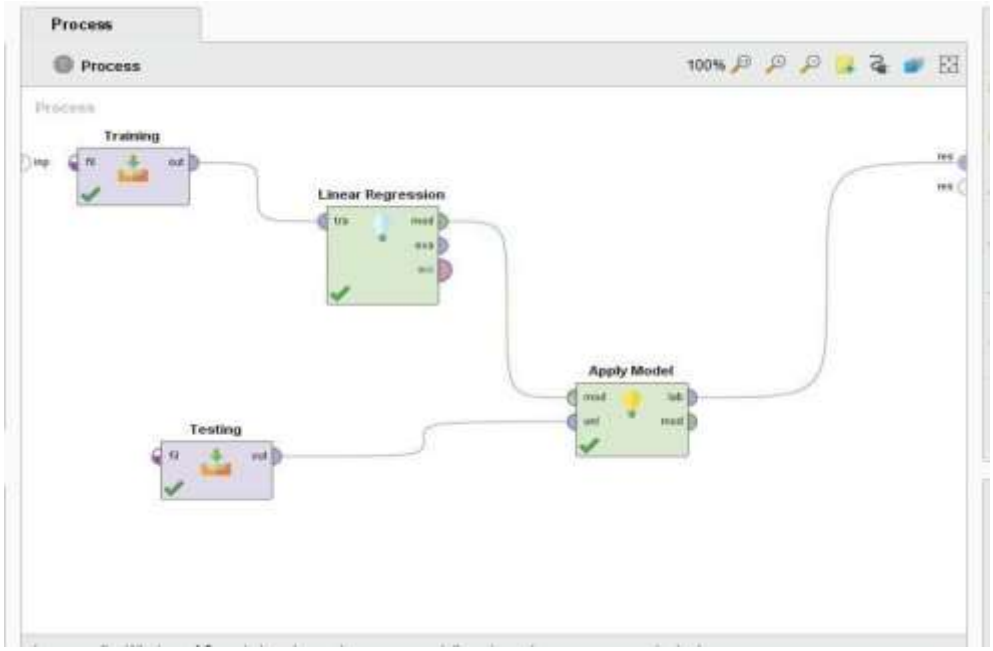

Figure 6. Prediction Data Model of Linear Regression Method

In the prediction process shown in Figure 6, there are two (2) additional process stages, namely:

a. Data Input Excel (Training and Testing)
   The data input process is used to enter data in Excel format into the process model.
b. Linear Regressions
   In the linear regression process, the data will be processed using a linear regression process.
c. Apply models.
   It is the final stage in which this process will apply a linear regression data mining model based on previous methods, starting from data training, data testing, and linear regression.

After all the processes are carried out, the prediction model results will be obtained, as shown in Figure 7.

Open in Turbo Prep Auto Model Filter (12 / 12 examples): all

Row No.	Bulan	Waktu Sandar	prediction(Waktu Sandar)	Bendera	Volume
1	Januari	118	120.741	16	3432
2	Februari	97	98.102	13	4423
3	Maret	92	90.555	12	4322
4	April	99	98.102	13	6544
5	Mei	98	98.102	13	7898
6	Juni	82	83.009	11	6445
7	Juli	99	98.102	13	2346
8	Agustus	96	98.102	13	9086
9	September	144	143.379	19	4344
10	Oktober	127	128.287	17	7764
11	November	152	150.926	20	2233
12	Desember	144	143.379	19	7445

Figure 7. Prediction Model Results

Using performance

Performance is used to measure the results of the model that has been formed and whether it is feasible to do it. Implementation in this study uses three (3) forms: RSME (root mean square error), absolute error, and relative error. Figure 8 is modeling using performance.

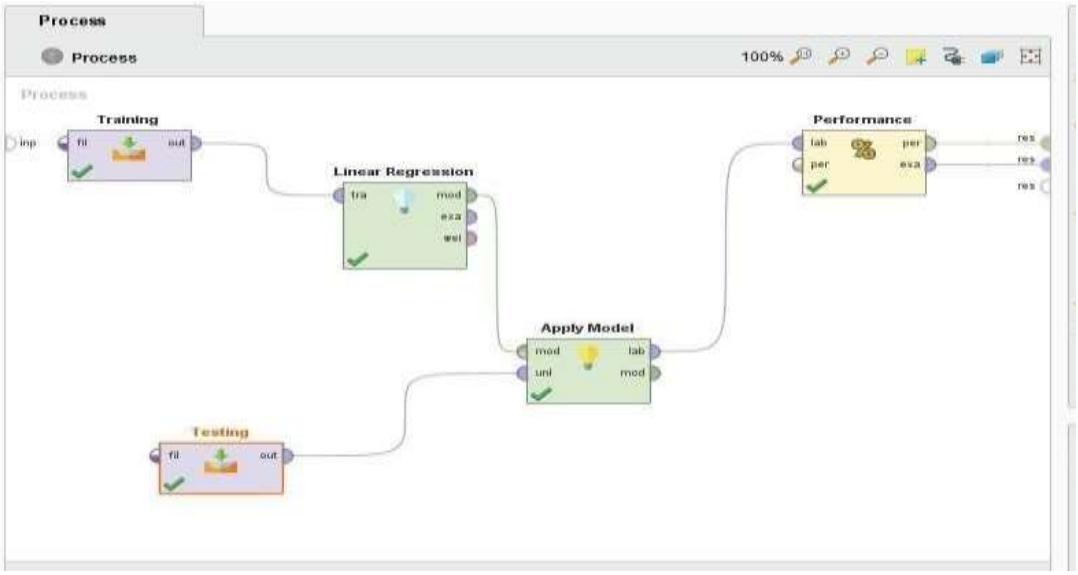

Figure 8. Process of Using Performance

In the prediction process for performance, there are several differences from the previous prediction process due to the addition of the performance operator (regression), where the purpose of the operator is to measure the root mean square error, absolute error, and relative error.

Multiple Linear Regression Model Testing

At this stage, testing is carried out for the prediction process of the following year's berthing time, namely 2022. Training data is taken from all data, while data testing is data taken for 2022, where the berthing time is unknown. Figure 9 is a prediction model.

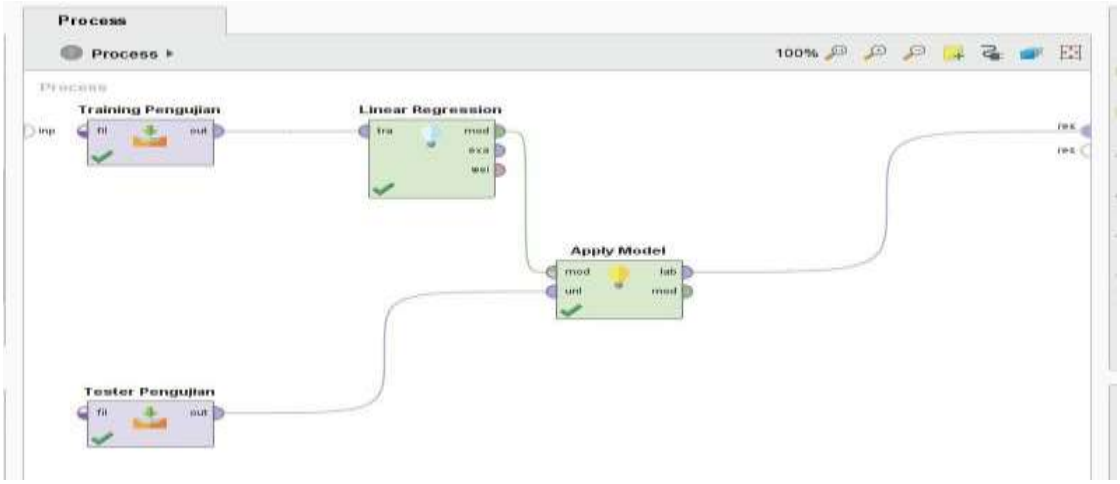

Figure 9. Prediction Model

In the prediction process, the prediction results are obtained as shown in Figure 10 below:

Open in Turbo Prep Auto Model Filter (12 / 12 examples): all

Row No.	Waktu Sandar	prediction(Waktu Sandar)	Bulan	Bendera	Volume
1	?	45.258	Januari	6	9888
2	?	45.258	Februari	6	2357
3	?	82.972	Maret	11	5423
4	?	60.344	April	8	2344
5	?	52.801	Mei	7	8764
6	?	45.258	Juni	6	2234
7	?	52.801	Juli	7	52
8	?	37.715	Agustus	5	55
9	?	52.801	September	7	4344
10	?	98.058	Oktober	13	3335
11	?	82.972	November	11	2393
12	?	90.515	Desember	12	7342

ExampleSet (12 examples, 2 special attributes, 3 regular attributes)

Figure 10. Prediction Results

The following is an explanation of the prediction results that have been carried out using the Rapidminer application:

Discussion of Prediction Models
In the prediction process with the rapid miner application, several results are displayed, including:

Table views

a. The intercept value obtained has a coefficient of 7,546.
b. The std error managing the value of null is 1095.
c. t-Star of 6,894.
d. The p-value of 0.000, generally in multiple linear regression to determine a significant value, namely:
    - If the p-value exceeds 0.05, the independent variable does not affect the dependent variable.
    - If the p-value is less than 0.05, the independent variable can influence the dependent variable.

In the prediction results that have been obtained, the resulting p-value is 0.000, which means the p-value is smaller than 0.05. Therefore, it can be concluded that the independent variable in this prediction affects the dependent variable. The table view display is shown in Figure 11 below.

Coefficient	Std. Error	Std. Coefficient	Tolerance	t-Stat	p-Value	Code
7.546	1.095	0.470	?	6.894	0.000	****

Figure 11. Table View Prediction

Data Views

The displayed data view table contains the predicted results of the ship's berthing time in 2021. In the data view shown by Figure 4.10, there are six (6) columns or variables displayed, including:

a. Row variable no is a variable that shows the serial number of the data. In row no variables, there are 1 to 12 that are displayed; this is because the process of predicting data input in data testing is from month 1 to month 12.
b. The lean time variable is a variable that shows the value of the slight time used for data testing in the prediction process. Rest times are shown by month.
c. Prediction variable: the prediction variable is a variable that shows the results of the prediction of the learning time that is generated from this prediction process. The results of the projection of lean times are displayed by month.
d. The flag variable is a variable that shows ships by flag in 2021. The total flags displayed are by month.
e. The volume variable is a variable that shows the total volume value in 2021. The total volume displayed is by month.

Meta Data View (Statistics)

The metadata view Contains information about the attribute table in the test data used to predict, such as name, type, statistics, range and missing. Meta data view on rapid miner is shown in Figure 12 below:

Figure 12. Meta Data View

Plot Views

Prior settings must be made to display the predicted results in the plot view. The following indicates the grounds for showing the prediction results in the plot view, as shown in Figure 13 below.

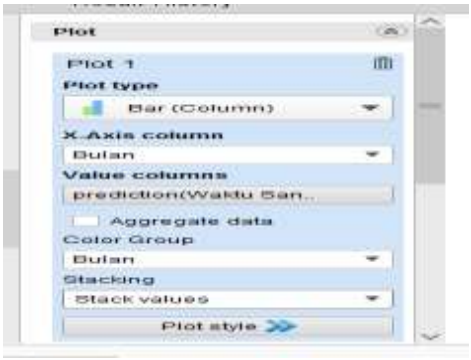

Figure 13. Plot View Settings

In the Plot view display setting, the plot type used is the bar (column); the bar is chosen because it can produce a diagram that illustrates the comparison of the total lean time value before prediction with the total slight time value after prediction is made. The year variable is used on the x-axis column to display the plot. In value columns, two (2) variables are used, namely leaning time (Y) and predicted leaning time (Y). Then, for stacking, no stacking is selected. After the settings are complete, the prediction results will appear as a diagram. The following is a display in Figure 14, and it is of the predicted learning time plot view in 2021:

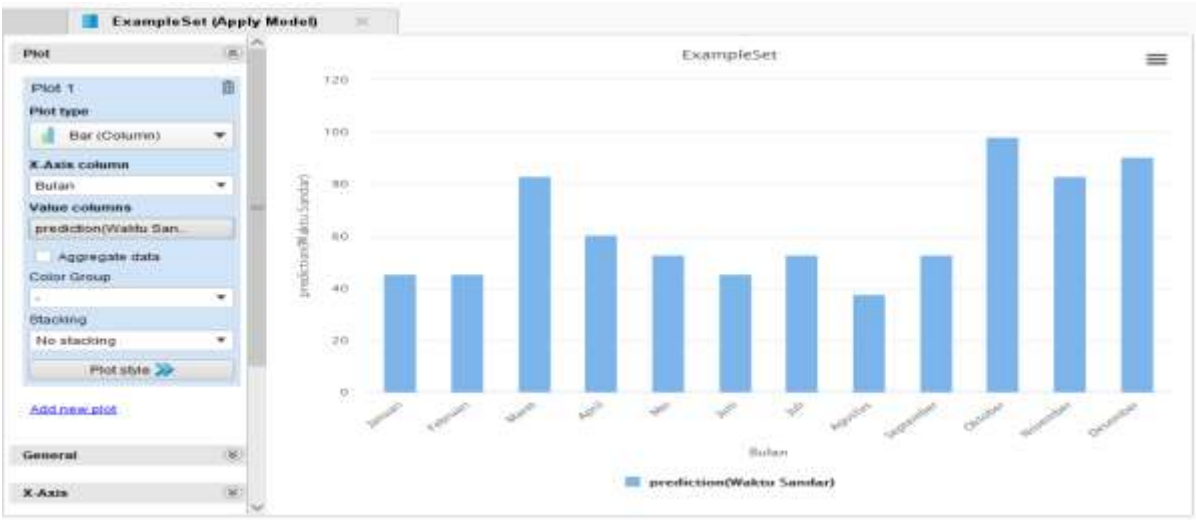

Figure 14. Plot Views

From the graph displayed in the plot view, the diagram shows information on lean times for each month in 2021.

Discussion of Performance

The results obtained from the prediction process to measure the error rate include the following results: the results obtained for the predictions to estimate the error rate using the 2021 testing data produce predictions, which can be seen in Figure 13—the results of the Prediction Testing data below, as shown in Figure 15.

prediction(Waktu Sandar)	Bulan
45.258	Januari
45.258	Februari
92.972	Maret
60.344	April
52.801	Mei
45.258	Juni
52.801	Juli
37.715	Agustus
52.801	September
98.058	Oktober
82.972	November
90.515	Desember

Figure 15. Discussion of Prediction Results of Ship Berthing Time in 2021

The test value of the model obtained a root mean square error (RSME) of 1.189, an absolute error of 1.075, and a relative error of 1.01%. Of all performance test values, all numbers have small values and are close to 0, so the modeling produces good predictions.

Conclusion

We have developed a model that accurately forecasts when the ship will berth. With prediction skills very similar to the original value, we presented this model at the New Boom Port of Palembang City. A total lean time of 573,263 is achieved for the year 2021 forecast results. Performance values yield a very good absolute error of 1.075, a relative error of 1.01%, and a root mean square error of 1.132. Because the findings display a very narrow error margin of less than 1.5, they are categorized as a very satisfactory model.

References

Akanmu, S., & Jaja, S. (2012). *Knowledge Discovery in Database: A knowledge management strategic approach*.

Hariaji, J. (2021). *Simple Linear Regression (SLR) Model and Multiple Linear Regression (MLR) Model*. https://doi.org/10.13140/RG.2.2.17237.35044

Kusrini, & Luthfi, E. T. (2009). *Algoritma Data Mining*. Andi Offset.

Laim, B., & Kumar, R. (2023). Data Mining Techniques for Weather Forecasting. *The Visual Computer*, *15*, 56–63.

Lenzerini, M. (2002). Data Integration: A Theoretical Perspective. In *Proceedings of the ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*. https://doi.org/10.1145/543613.543644

Muzakir, A., & Wulandari, R. A. (2016). Model Data Mining sebagai Prediksi Penyakit Hipertensi Kehamilan dengan Teknik Decision Tree. *Scientific Journal of Informatics*, *3*(1). http://journal.unnes.ac.id/nju/index.php/sji

Oliveira, E., Tostes, M., Freitas, C., & Leite, J. (2017). Voltage THD Analysis Using Knowledge Discovery in Databases With a Decision Tree Classifier. *IEEE Access*, *PP*, 1. https://doi.org/10.1109/ACCESS.2017.2778028

Luh, N., Wulandari, P., Kartika, A., Sarja, Y., Ayu, G., & Saryanti, D. (2015). PREDIKSI JUMLAH PELANGGAN DAN PERSEDIAAN BARANG MENGGUNAKAN METODE REGRESI LINIER BERGANDA PADA BALI ORCHID. *Eksplora Informatika*, *L2*, 1–12.

Rahman, F., Kassim, R., Baharum, Z., Noor, H., & Haris, N. (2019). *Data Cleaning in Knowledge Discovery Database-Data Mining (KDD-DM)*. https://doi.org/10.35940/ijeat.F1100.0986S319

Saut Gurning, R. O., & Budiyanto, E. H. (2007). *Manajemen Bisnis Pelabuhan*. APE Publishing.

Triadmodjo, B. (2009). *Perencanaan Pelabuhan*. Beta Offset, Jogyakarta.