DETERMINATION PRICE TICKET OF AIRLINE LOW-COST CARRIER BASED ON DYNAMIC PRICING STRATEGY USING MULTIPLE REGRESSION METHOD

PENENTUAN HARGA TIKET AIRLINE LOW-COST CARRIER BERDASARKAN STRATEGI DYNAMIC PRICING MENGGUNAKAN METODE LINEAR BERGANDA

Elisa Intan Puspita Sari [1], Rio Aurachman[2], Mohammad Deni Akbar[3]

[1,2,3] Universitas Telkom, Bandung
[1]**elisaintanp@student.telkomuniversity.ac.id**, [2]**rioaurachman@telkomuniversity.ac.id**, [3]**denimath@telkomuniversity.ac.id**

Abstract

To maintain the company's position in the aviation industry, each airline has its strategy to overcome losses due to the impact of the Covid-19 pandemic. One strategy that is often applied by airlines is the dynamic pricing strategy. The ever-changing flight ticket prices serve to maximize the company's revenue. The puIDRose of this study is to propose the optimum pricing with a dynamic pricing model which in this model will produce an optimum pricing policy based on historical sales. This final project uses dynamic pricing to maximize revenue by modeling the effect of price on demand. The method used in this research is Multiple Regression with the help of Jupyter Notebook software in Python in the data processing. Furthermore, the author uses cross-validation which serves to evaluate the demand model and price prediction used. This final project using data from PT. Trigana Air with flights from Wamena-Jayapura on 01 February 2021 to 28 February 2021. The demand model used produces an accuracy of 61.24% with a standard deviation of 28,6%. Meanwhile, the ticket price prediction model used produces an accuracy of 55,4% with a standard deviation of 22,5%. Based on the dynamic pricing model that has been applied by optimizing sales profits based on the optimal price variable, there is an increase in revenue on March 12, 2021 with a profit of IDR38.220.644 or an increase of 22,6%.

Keywords: multiple regression, machine learning, dynamic pricing

Abstrak

Dalam rangka mempertahankan posisi perusahaan di industri penerbangan, setiap maskapai memiliki strategi tersendiri untuk mengatasi kerugian akibat dampak pandemi Covid-19. Salah satu strategi yang sering diterapkan oleh maskapai penerbangan adalah penentuan harga tiket pesawat secara dinamis atau dynamic pricing strateegy. Harga tiket pesawat yang selalu berubah-ubah berfungsi untuk memaksimalkan pendapatan perusahaan. Tujuan dari penelitian ini adalah untuk mengusulkan penetapan harga optimum dengan model dynamic pricing dimana pada model ini akan menghasilkan kebijakan harga yang optimum berdasarkan penjualan historis. Tugas Akhir ini menggunakan dynamic pricing untuk memaksimalkan pendapatan dengan memodelkan pengaruh harga terhadap permintaan. Metode yang digunakan dalam penelitian ini adalah Multiple Regression dengan bantuan software Jupyter Notebook berbahasa Python dalam pengolahan data. Selanjutnya penulis menggunakan cross-validation yang berfungsi untuk mengevaluasi model permintaan dan prediksi harga yang digunakan. Dalam tugas akhir ini menggunakan data PT. Trigana Air dengan penerbangan dari Wamena-Jayapura pada tanggal 01 Februari 2021 sampai tanggal 28 Februari 2021. Model permintaan yang digunakan menghasilkan akurasi sebesar 61,4% dengan standar deviasi sebesar 28,6%. Sedangkan untuk model prediksi harga tiket yang digunakan menghasilkan akurasi sebesar 55,4% dengan standar deviasi 22,5%. Berdasarkan model dynamic pricing yang telah diterapkan dengan mengoptimalkan keuntungan penjualan berdasarkan variabel harga yang optimal didapatkan peningkatan keuntungan penjualan pada tangal 12 Maret 2021 dengan keuntungan seesar Rp 38.220.644 atau mengalami kenaikan 22,6%.

Kata kunci: *regresi linear berganda*, machine learning, dynamic pricing

I. Prelimary

The airline industry in Indonesia is experiencing rapid development, where according to International Air Transportation (IATA) predicts that the frequency of flights in Indonesia in 2020 will be included in the top 10 in the world. Based on data from the Directorate General of Civil Aviation (DGCA), There are 12 airlines operating and registered in Indonesia. However, of the 12 airlines, only 8 airlines control and serve strategic routes and the others serve short routes and become beginner or pioneer aircraft for remote areas of Indonesia.

In early 2020, many countries decided to close or reduce the frequency of international and domestic flights to suppress the spread of the coronavirus and cause a decline in the aviation industry. With the drastic reduction in the frequency of scheduled flights or scheduled light from March to May 2020 without a clear time, the aviation industry has suffered a lot of losses. This is also felt in Indonesia. On may 7, 2020, the Minister of Transportation changed the rule, whereby all modes of transportation may operate again to transport passengers, including airlines, with restrictions and conditions that must be met, such as having a cover letter from the local area, a health letter, or a Covid-19 free certificate, and must implement national health protocols and standards.

The policy of allowing passenger airlines to operate during the covid-19 pandemic has not been able to make the aviation industry able to overcome its losses, so the Ministry of Transportation decided to allow airlines to increase the price of airline tickets during the covid-19 pandemic following the upper limit tariffs that have been set in the Decree. Minister Number 106 of 2019 concerning Tariffs for the Upper Limit of Domestic Economic Service Passengers of Scheduled Commercial Air Transport.

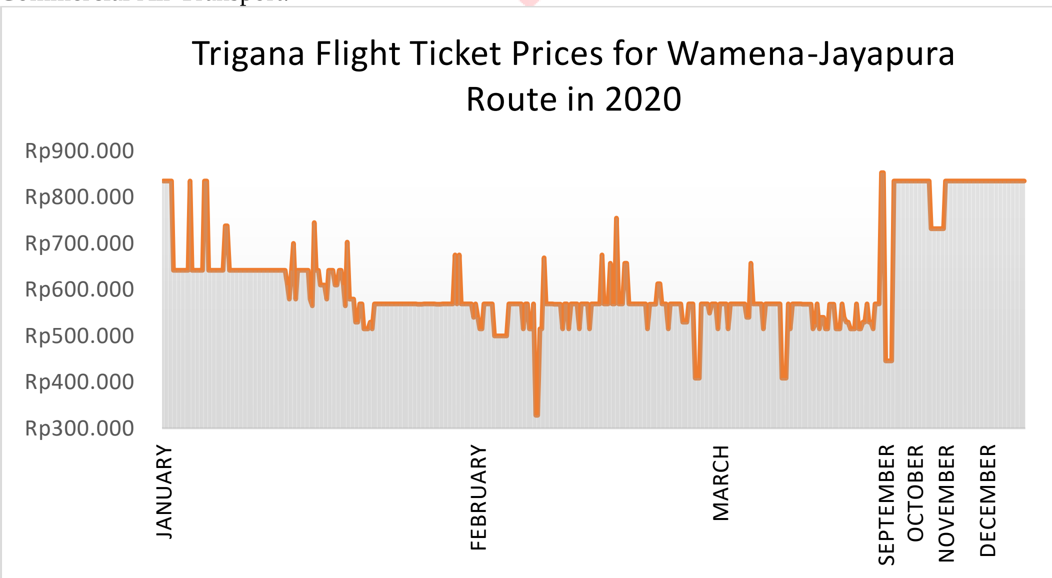


Figure I. 1 Graph of Wamena-Jayapura Airline Ticket Prices in 2020
Source (PT. Yunejer Travel Cakrawala Wisata, 2021)

Figure I. 1 shows Trigana airline ticket prices for Wamena-Jayapura flights in 2020. The data presented is sales data for 7 months, where from April to August there are flight restrictions or airport access closures so for 5 months there is no flight activity. And during the sales there are the lowest price of IDR. 328,500/pax and the highest price is IDR. 835,000/pax. In the conditions of the Covid-19 pandemic, the airline industry has made many changes to airline ticket prices to maintain the company's position in competition in the aviation industry. Every airline has a strategy to survive in the competition between airlines. One of the strategies implemented by airlines is dynamic pricing of airline tickets or those that are adjusted to the uncertainty of demand.

Dynamic Pricing is a dynamic pricing strategy within the Airlines Revenue Manager in the form of pricing arrangements that aim to increase airline revenue. Dynamic Pricing is often applied in the aviation industry because it has characteristics that match the pricing of airline tickets which are determined based on certain criteria. These criteria can be in the form of remaining sales time, remaining unsold seats, flight times or schedules, competitor prices, and others. The advantage of implementing a dynamic pricing strategy is that the company or airline can increase or decrease prices according to the company's desire to increase or earn revenue.

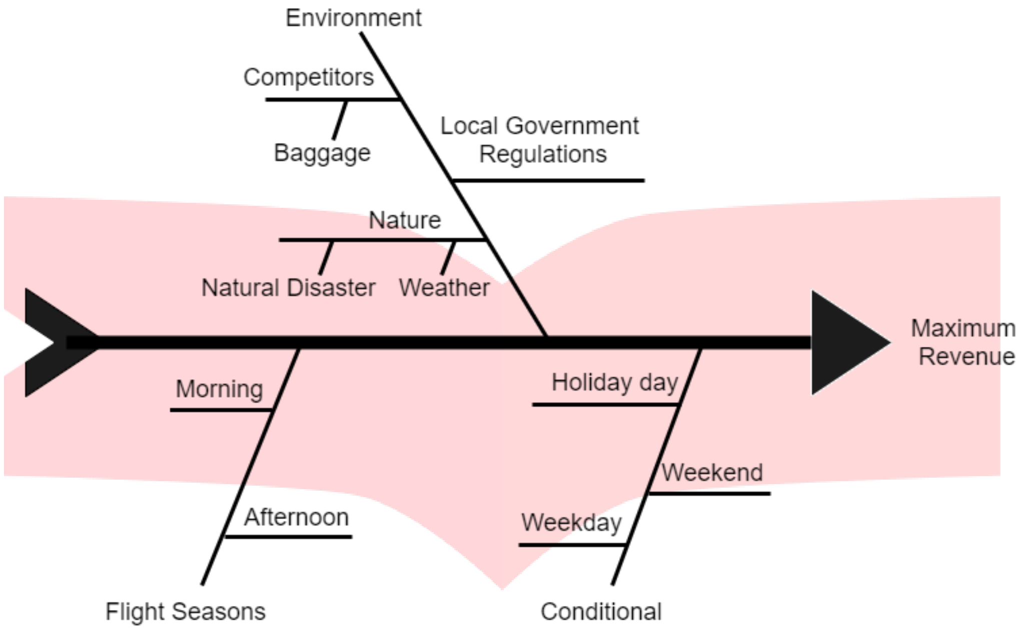


Figure I. 2 Fishbone

Figure I. 2 shows that maximizing airline revenue is influenced by several factors such as conditional factors where on national holidays or weekdays and weekends there are different price variations, then there are flight departure time factors such as flight times in the morning and evening which affect demand, and factors environment such as competitors, nature and regional regulations during the covid-19 pandemic where in order to overcome the spread of covid-19 many regional regulations prohibit flights for a certain period, this results in airlines experiencing losses that affect revenue.

Steps that can be taken to overcome these problems are by making a daily flight ticket price prediction system. One method that can be used to predict prices with continuous numerical value is by using the regression method. Regression aims to find a function that models the data. In creating and processing predictive models, the author uses machine learning tools with the aim of getting results that are more efficient, accurate, and can shorten time.

Machine learning is one technology that can provide solutions and convenience in designing aircraft price prediction models, where the application of machine learning techniques can make it easier for companies to deal with future price uncertainty by using data from the previous period.

The prediction process in machine learning is to understand the nature or characteristics of unknown objects by identifying patterns in the dataset. The hallmark of machine learning is the existence of training, learning, or training. Therefore, machine learning requires data to be studied as training data (training set), in this case, the variables that affect the price of airline tickets such as the distance to order airline tickets, departure schedules, holidays, and others. The model that will be generated from the training process will be used as a reference in predicting the price of daily air tickets.

Based on this description, in this final project, a strategy for changing flight ticket prices from airlines for short routes is carried out by processing data using machine learning to maximize revenue. This study was conducted to determine the maximum revenue based on the demand and ticket price prediction model.

II. Literature Review

II.1 Machine Learning

Machine Learning (ML) is one of the branches of Artificial Intelligence, especially those that study how computers can learn from data to improve their intelligence [1]. The characteristics of Machine Learning is the existence of a training, learning, or training process. So that ML needs data to be learned or what is commonly called training data. Several processes to build a machine learning system, namely collecting data, preparing input data, analyzing input data, involving human involvement, training algorithms, testing algorithms, using algorithms[2].

II.2 Cross Validation

Cross Validation (CV) is a method for estimating prediction error for evaluating model performance. In cross-validation known as rotation estimation, by dividing the data into k subsets of almost equal size, the model in the classification is trained and tested for-k. In each iteration, one subset will be used as test data and the other k data subgroups will function as training data[3]. K-fold cross-validation is a method for evaluating classifier performance, this method can be used if it has a limited amount of data (the number of instances is not much). K-fold cross-validation is a method used to determine the average success of a system by performing redundancy by randomizing the input attributes so that the system is tested for several random input attributes.

II.3 Low-Cost Carrier

The low-cost carrier is a redefinition of the airline business by striving for the most efficient price possible to reduce operational costs so that it can meet all market segments by reducing various minimalist facilities and services.

II.4 Revenue Management

Revenue management is a systematic approach to implementing pricing and inventory controls on the sale of a perishable asset. Airline supply is limited by aircraft capacity and is highly perishable, where unsold seats cannot be reused after the flight departs. Under these conditions, the process of pricing and inventory control in the aviation industry is a complex matter that industry players need to deal with[4].

II.5 Price

Price is one element of the marketing mix or marketing mix that can generate revenue, where other elements get costs[5]. Price is the amount of money needed to obtain several combinations of a product and service that accompanies it. The definition of price can be interpreted as the amount of money used to assess and obtain a good or service.

II.6 Dynamic Pricing Model

Dynamic Pricing is a dynamic pricing strategy within the Airlines Revenue Manager in the form of pricing arrangements that aim to increase airline revenue. Dynamic pricing models the effect of product prices at different times from the demand model (Shakya, Kern, Owusu, & Chin, 2012). The dynamic pricing model is explained by equation II-1 below:

$$\pi = \sum_{t=1}^{N} (P_t \, Q_t) \quad \text{(II-1)}$$

Where:

\Pi = Total revenue during the planning horizon
N = Number of periods in the planning horizon
t = Period t on the planning horizon
Q_t = Number of demands in period t
P_t = The predict price of the product in period t

II.7 Multiple Regression

Multiple regression is used to determine the proportion of a variable that is continuous and related. The general model of linear multiple in the population can be seen in equation II-2 below[6]:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \varepsilon \quad \text{(II-2)}$$

Where:

$\boldsymbol{\beta_0}$ $= \mathbf{The\ Y - intercept}$
$\boldsymbol{\beta_k}$ = The slope of the regression surface with respect to the variable $\mathbf{X_k}$
$\boldsymbol{\varepsilon}$ $=$ **Error term**

II.8 Asumptions for Multiple Regression

The classical assumption is one of the tests used for statistical requirements. The classical assumption test of the linear regression model used is carried out to know whether the regression model is good or not. The purpose of classical assumption testing is to provide certainty that the regression equation obtained is accurate in estimation, unbiased, and consistent. The assumptions that must be met in the regression analysis are normality, homoscedasticity, multicollinearity, and autocorrelation[8].

II.9 Hypothesis Testing

Hypothesis testing was conducted to determine whether the parameters obtained from the estimation were statistically significant. There are two types of hypothesis testing in the multiple regression model, namely individual hypothesis testing and multiple hypothesis testing. Alpha (α) used in testing this hypothesis is 0.05. The initial hypothesis for this test is:
H0: There is no significant effect of the independent variable on the dependent variable.
H1: There is a significant effect of the independent variable on the dependent variable.

III. Problem Solving Methodology

III.1 Preliminary Phase

At this phase, a preliminary study of the object of research is carried out to facilitate the problems to be raised in this research. First, conduct a literature study by reviewing previous studies related to various sources. Then the problem will be formulated by systematically formulating existing conditions based on theories regarding the cause and effect of these problems. Then from the results of the problem formulation, the research objectives were determined with predetermined limitations.

III.2 Data Collection Phase

At this phase, the author uses primary data and quantitative secondary data as material for the analysis and design of a plane ticket price prediction model. Primary data includes daily sales data for the airline Trigana Air for one week in February 2021 and respondent data from a comparative questionnaire for comparison of existing and newly created systems. While secondary data in the form of supporting statistical data taken from websites such as the Central Statistics Agency (www.bps.com) in the form of international and national flight traffic data in 2015-2019.

III.3 Model Development Phase

At this phase, multiple regression models will be developed. The development of this model is carried out based on the data that has been collected by determining the variables that will be used to develop a model based on each flight schedule. The development of this model also refers to the regression model in the reference book " Business Applications of Multiple Regression ". In processing multiple regression models with machine learning, the data will be divided into 70% training data and 30% testing data. After forming the regression model, an experiment was carried out in estimating the regression model parameters such us intercept value and coefficient independent variables values.

III.4 Evaluation Method Phase

At this stage, validate the prediction model using 10-cross validation to measure the level of accuracy and standard deviation of the ticket price prediction model. And the last is hypothesis test will be conducted to analyze the relationship between variables from the model that has been made.

IV. Discussion

IV.1 Regression Model Development

Process of learning the model on the data set. The estimator used is linear regression. Here in Figure 7 the learning output of the linear regression model.

```
Out[9]: LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None,
                 normalize=False)
```

Figure IV. 1 Machine Learning Output of Linear Regression Model

Figure IV. 1 shows the output message after the linear regression model learning process on the data set (X and Y).

After the machine learning process of the linear regression model, the intercept values and linear regression coefficients are obtained as follows:

Table IV. 1 Intercept and Coefficient Demand Prediction

Intercept	163.05995219373386
Coefficient	
X1	19.983361
X2	21.036416
X3	14.877657
X4	-4.969687
X5	3.256999

Based on Table IV. 1, the results of the intercept and coefficient values of each independent variable which can show the relationship between the dependent variable and the independent variable based on the direction. The variables X4 and X6 have negative results (-) which means when the independent variable increases, the dependent variable will decrease. As for the variables X1, X2, X3 and X5 have positive results (+) which means when the independent variable increases, the dependent variable will also increase. Based on the results of multiple linear regression learning in machine learning, the regression model is as follows:

$$Y = 163.05995219373386 + 19.983361X_1 + 21.036416X_2 + 14.877657X_3 - 4.969687X_4 + 3.256999X_5 - 0.000199X_6 + \varepsilon \quad \text{(IV-1)}$$

Where:

Y = Dependent variable (demand forecast)

$X_1, X_2, X_3, X_4, X_5, X_6$ = *Independent variables*

X_1 = Firt flight variable

X_2 = Second flight variable

X_3 = Holiday day variable

X_4 = Day of week variable

X_5 = Baggage facility variable

X_6 = Price in the previous period variable

ε = Error term

Table IV. 2 Intercept and Coefficient Price Prediction

Intercept	182759.10942718107
Coefficient	
X1	6358.381111
X2	91874.352634
X3	78097.678975
X4	-38236.527887
X5	36735.138679

Based on Table IV. 2, the results of the intercept and coefficient values of each independent variable which can show the relationship between the dependent variable and the independent variable based on the direction. The variables X2 and X3 has negative results (-) which means when the independent variable increases, the dependent variable will decrease. As for the variables X1, X4 and X5 have positive results (+) which means when the independent variable increases, the dependent variable will also increase. Based on the results of multiple linear regression learning in machine learning, the regression model is as follows:

$$Y = 182759.10942718107 + 6358.381111X_1 + 91874.352634X_2 + 78097.678975X_3 - 38236.527887X_4 + 36735.138679X_5 + 203642.021227X_6 + \varepsilon \text{ .... (IV-2)}$$

Where:

Y = Dependent variable (price in the previous period)

$X_1, X_2, X_3, X_4, X_5, X_6$ = Independent variables

X_1 = Sales ticket in the previous period variable

X_2 = Fisrt flight variable

X_3 = Second flight variable

X_4 = Holiday day variable

X_5 = Day of week variable

X_6 = Baggage facility Variable

ε = Error term

IV.2 Data Procssing

IV.2.1 Normality

A normality test is a test carried out intending to assess the distribution of data on a normally distributed variable or not. To test whether the data is normally distributed or not, the Kolmogorov-Smirnov test was performed statistically. The residual is normally distributed if it has a significance value > 0.05. The results of the normality assumption test using Kolmogorov-Smirnov are as follows:

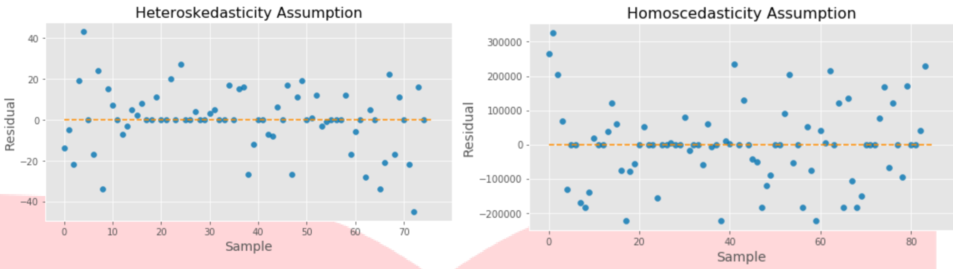
Table IV. 3 Normality Test Result

P-value from Test Kolmogorov Smirnov	
Demand Prediction Model	1.360600715008906e-07
Price Prediction Model	2.1431837409628652e-05

Based on Table IV. 3, the results of the normality test using the Kolmogorov-Smirnov for the demand prediction model has a p-value of 1,36 and the price prediction model has a p-value of 2,14. so it can be concluded that the p-value of the demand and price prediction model is more than 0.05, which means that the demand and price prediction model is normally distributed.

IV.2.2 Heteroskedasticity

The variance of the error is homogeneous (homoskedastic), that is, the error has the same variance value. In this final project the author uses a scatterplot graph to test heteroscedasticity.



(a)Demand Prediction (b) Price Prediction

Figure IV. 2 Heteroskedasticity Test Result Demand Prediction

Figure IV. 2 shows that from the results of the Scatter Graph there is no certain pattern formed because the points spread irregularly above and below the 0 axis on the Y axis. It can be concluded that there are no symptoms of heteroscedasticity.

IV.2.3 Multicollinearity

Multicollinearity means that there is a close linear correlation between the independent variables. The statistical test used is the Variance Inflation Factor (VIF). The Variance Inflation Factor (VIF) value greater than 10 indicates serious multicollinearity. The results of the multicollinearity assumption test using VIF statistics are as follows:

Table IV. 4 Multicollinearity Test Result

Demand Prediction Model			Price Prediction Model		
	Features	VIF		Features	VIF
5	X6	5.13	0	X1	9.55
4	X5	3.14	2	X3	5.29
1	X2	2.47	1	X2	4.82
0	X1	2.36	5	X6	2.81
3	X4	1.57	4	X5	1.38
2	X3	1.11	3	X4	1.21

Based on Table IV. 4, the value of the variance inflation factor (VIF) of the five independent variables is less than 10, this indicates that there is no multicollinearity among the independent variables.

IV.2.4 Autocorrelation

One way to detect the presence or absence of autocorrelation in a regression model is to test the Durbin-Watson test value with the following conditions [9]:

1. There is a positive autocorrelation if the DW value is below lower bound (dL) or DW < dL
2. There is no autocorrelation if the DW value is between upper bound (dU) and (4 – dU) or dU < DW<(4 – dU)
3. There is no negative autocorrelation if the DW value is above (4 – dL) or DW > (4 – dL)

Table IV. 5 Autocorrelation Test Result

	Durbin-Watson
Demand Prediction Model	1.9659462342638825
Price Prediction Model	2.193967071265473

Based on Table IV. 5, the Durbin-Watson value for demand prediction model is 1,9659 and Durbin-Watson value for price prediction model is 2,1939. It can be concluded that Durbin-Watson value of the demand prediction still in the range of 1,7698 < 1,9659 < 2,2302 and as well as Durbin-Watson value of price prediction model still in the range of 1,7698 < 2,1939< 2,2302 so that it identifies no autocorrelation.

IV.3 Analysis and Validation of Implementation System

IV.3.1 Cross Validation Testing

In making a demand and price prediction model using machine learning, the accuracy and standard deviation for each model is calculated using the cross-validation method, so that the results are as below:

Table IV. 6 Accuracy and Standard Deviation Price Prediction Model

	Accuracy	Standard Deviation
Demand Prediction Model	0.6144668736935588	0.2860196836615051
Price Prediction Model	0.5542446307168835	0.225168933336370533

Based on Table IV. 6, the results of the calculation of the average error metric from 10-cross validation for demand prediction model is 61,4% with a standard deviation is 28,6%. Meanwhile the average error metric for price prediction model is 55,4% with a standard deviation is 22,5%. The standard deviation of the cross-validation results works to see the distance between the average accuracy of each training (iteration).

After re-predicting using cross-validation, the results of the OLS summary from the linear regression model are as follows:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      Y   R-squared:                       0.562
Model:                            OLS   Adj. R-squared:                  0.510
Method:                 Least Squares   F-statistic:                 1.273e+21
Date:                Sat, 14 Aug 2021   Prob (F-statistic):               0.00
Time:                        16:49:44   Log-Likelihood:                -219.87
No. Observations:                  58   AIC:                             453.7
Df Residuals:                      51   BIC:                             468.2
Df Model:                           6
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const        163.0600     88.988      1.832      0.000     -15.591     341.711
X1            19.9834      5.134      3.892      0.000       9.676      30.291
X2            21.0364      4.998      4.209      0.000      11.002      31.071
X3            14.8777      8.763      1.698      0.000      -2.715      32.470
X4            -4.9697      3.199     -1.553      0.000     -11.392       1.453
X5             3.2570      4.403      0.740      0.000      -5.582      12.096
X6            -0.0002      0.000     -1.600      0.000      -0.000    5.08e-05
==============================================================================
Omnibus:                        2.581   Durbin-Watson:                   1.665
Prob(Omnibus):                  0.275   Jarque-Bera (JB):                1.712
Skew:                          -0.291   Prob(JB):                        0.425
Kurtosis:                       3.607   Cond. No.                     4.22e+07
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

Figure IV. 3 OLS Summary Demand Prediction Model

Figure IV. 3 shows the summary results of the regression model that has been made. Based on the picture, the R-squared is worth 0,562 which means 56,2% of the influence of the independent variable on the dependent variable.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      Y   R-squared:                       0.684
Model:                            OLS   Adj. R-squared:                  0.647
Method:                 Least Squares   F-statistic:                     18.39
Date:                Sat, 14 Aug 2021   Prob (F-statistic):           3.09e-11
Time:                        17:18:00   Log-Likelihood:                   0.00
No. Observations:                  58   AIC:                             1551.
Df Residuals:                      51   BIC:                             1565.
Df Model:                           6
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const       1.828e+05   4.77e+04      3.831      0.000     8.7e+04    2.79e+05
X1          6358.3811   1609.357      3.951      0.000    3127.462    9589.301
X2          9.187e+04   7.21e+04      1.274      0.000   -5.29e+04    2.37e+05
X3           7.81e+04    7.2e+04      1.085      0.000   -6.65e+04    2.23e+05
X4         -3.824e+04   1.14e+05     -0.336      0.000   -2.67e+05    1.91e+05
X5          3.674e+04    4.2e+04      0.875      0.000   -4.76e+04    1.21e+05
X6          2.036e+05   4.23e+04      4.810      0.000    1.19e+05    2.89e+05
==============================================================================
Omnibus:                        3.861   Durbin-Watson:                   1.991
Prob(Omnibus):                  0.145   Jarque-Bera (JB):                2.416
Skew:                           0.294   Prob(JB):                        0.299
Kurtosis:                       2.191   Cond. No.                         257.
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

Figure IV. 4 OLS Summary Price Prediction Model

Figure IV. 4 shows the summary results of the regression model that has been made. Based on the picture, , the R-squared is worth 0,684 which means 68,4% of the influence of the independent variable on the dependent variable.

IV.3.2 Individual Hypothesis Test

Based on Figure IV. 3 and Figure IV. 4, it can be seen the relationship between the independent and dependent variables from the p-value (P>|t|) as follows:

1. Variable X1 has a significance value of t(0.00) <0.05, so there is a significant effect between the independent variables on the dependent variable. This means that Ho is rejected and accepts H1.
2. Variable X2 has a significance value of t(0.00) <0.05, so there is a significant effect between the independent variables and the dependent variable. This means that Ho is rejected and accepts H1.
3. Variable X3 has a significance value of t(0.00) <0.05, so there is a significant effect between the independent variables and the dependent variable. This means that Ho is rejected and accepts H1.
4. Variable X4 has a significance value of t (0.00) <0.05, so there is a significant effect between the independent variables and the dependent variable. This means that Ho is rejected and accepts H1.
5. Variable X5 has a significance value of t(0.00) <0.05, so there is a significant effect between the independent variables and the dependent variable. This means that Ho is rejected and accepts H1.
6. Variable X6 has a significance value of t (0.00) <0.05, so there is a significant effect between the independent variables on the dependent variable. This means that Ho is rejected and accepts H1.

IV.3.3 Multiple Hypothesis Testing

Based on Figure IV. 3, it can be seen the relationship between the independent and dependent variables from the F-count (F-statistic) value of 1,27 with a significance Prob(F-statistic) value of (0.00) <0.05, then the independent variable influences dependent variable which means rejecting H0 and accepting H1. Meanwhile based on Figure IV. 4, it can be seen the relationship between the independent and dependent variables from the F-count (F-statistic) value of 3,09 with a significance Prob(F-statistic) value of (0.00) <0.05, then the independent variable influences dependent variable which means rejecting H0 and accepting H1.

IV.3.4 Optimation Revenue

At this stage, calculations the maximum and minimum revenue using the data from the demand and price predictions obtained in the previous subchapter with implementation dynamic pricing model.

Table IV. 7 Optimum Revenue

Date	Flight	Demand Prediction	Price Prediction	Revenue	%Deviation
12/03/2021	2	62	IDR616.462	IDR38.220.644	22,6%
06/03/2021	1	44	IDR708.777	IDR31.186.188	0,0%
28/03/2021	3	16	IDR603.823	IDR9.661.168	-69,0%
Average		40	IDR636.659	IDR25.563.986	

Based on Figure IV. 7 it can be concluded that when demand is at the maximum and minimum point there is an effect on revenue. This can be seen from the results of the % Deviation where for the maximum demand on March 12, it produces a positive calculation which means there is an increase in revenue of 22,6% with a revenue of IDR38.220.644, while for the minimum demand on March 28, has a negative calculation result which means there is a decrease revenue of -69,0% with a revenue of IDR 9.661.168.

V. Conclussion

Based on the results of this final project, the following conclusions can be drawn:

1. The demand prediction model using the multiple regression method can be said to be good. This is based on the results of the coefficient of determination of 56,2% with the conclusion that each variable influences the other. The model requested has an accuracy rate of 61,4% with a standard deviation of 28,6% so that it can be said to be quite feasible in predicting demand.
2. The price prediction model using the multiple regression method can be said to be in a good category. This is based on the results of the coefficient of determination of 68,4% with the conclusion that each variable influences the other. The price prediction model has an accuracy rate of 55,4% with a standard deviation of 22,5% so that it can be said to be quite feasible in predicting demand.
3. Based on the Dynamic Pricing model used in optimizing the sales revenue based on the optimum revenue, a maximum revenue increase of 22,6% was obtained on March 12, 2021 on the second flight of IDR38.220.644 with 62 demand.

REFERENCES

[1] T. Wahyono, *Fundamental of Python for Machine Learning: Dasar-Dasar Pemrograman Python untuk Machine Learning dan Kecerdasan Buatan*. Yogyakarta: Gava Media, 2018.

[2] P. Harrington, *Machine Learning in Action*. Manning Pulications, 2012.

[3] Nurhayati, I. Soekarno, I. K. Hadihardaja, and M. Cahyono, “A study of hold-out and k-fold cross validation for accuracy of groundwater modeling in tidal lowland reclamation using extreme learning machine,” *IEEE 2nd Int. Conf. Technol. Informatics, Manag. Eng. Environ.*, pp. 228–233, 2014.

[4] A. Cento, *The Airline Industry*, Challenges. Heidelberg : Physica-Verl, 2009.

[5] P. Kotler and K. L. Keller, *Marketing Management*, Thirteenth. Jakarta: Erlangga, 2009.

[6] R. Richardson, *Business Applications of Multiple Regression*. 2011.

[7] A. Izzah, “Prediksi Harga Saham Menggunakan Improved Multiple Linear Regression untuk Pencegahan Data Outlier,” *Kinet. Game Technol. Inf. Syst. Comput. Netw. Comput. Electron. Control*, 2017, doi: 10.22219/kinetik.v2i3.268.

[8] I. Ghozali, *No TitleAplikasi Analisis Multivariate Dengan Program SPSS*. Semarang: Badan Penerbit Universitas Diponegoro, 2013.

[9] D. Sunyoto, *METODOLOGI PENELITIAN AKUNTANSI*. Bandung: Refika Aditama Anggota Ikapi, 2016.