Prediction of New Student Numbers using Least Square Method

Dwi Mulyani

College of Informatics And Computer Management (STMIK) Banjarbaru
Banjarbaru Kalsel, Indonesia

Abstract—**STMIK BANJARBARU has acquired less number of new students for the last three years compared to the previous years. The numbers of new student acquisition are not always the same every year. The unstable number of new student acquisition made the difficulty in designing classes, lecturers, and other charges. Knowing the prediction number of new student acquisition for the coming period is very important as a basis for further decision making. Least Square method as the method of calculation to determine the scores prediction is often used to have a prediction, because the calculation is more accurate then moving average.**

The study was aimed to help the private colleges or universities, especially STMIK BANJARBARU, in predicting the number of new students who are accepted, so it will be easier to make decisions in determining the next steps and estimating the financial matters.

The prediction of the number of new student acquisition will facilitates STMIK BANJARBARU to determine the number of classes, scheduling, etc.

From the results of the study, it can be concluded that prediction analysis by using Least Square Method can be used to predict the number of new students acquisition for the coming period based on the student data in the previous years, because it produces valid results or closer to the truth. From the test results in the last 3 years, the validity shows 97.8%, so it can be said valid.

Keywords—Prediction of New Students; Least Square method

I. INTRODUCTION

STMIK BANJARBARU as one of the colleges in the field of computer becomes one destination for new student candidates to continue their education. In the first year, the number of student candidates could be predicted since STMIK BANJARBARU was the only one computer college in South Kalimantan. However, in the last few years there are many other universities in South Kalimantan which provide department of computer science so it is assumed that the number of student candidates are divided into some computer universities in South Kalimantan. This causes the regression of the number of new students in the last three years. Based on the student data of STMIK BANJARBARU, the numbers of New Students accepted were 407 in 2009, 516 in 2010, 528 in 2011, 374 in 2012, 375 in 2013 and 386 in 2014. In 2012 to 2014, the number of new students accepted in STMIK BANJARBARU has decreased to 29.16%.

The problem faced by STMIK BANJARBARU is estimating the number of new students due to the regression of the number of new student acquisition in the last three years.

The study was aimed to help the private universities, especially STMIK BANJARBARU, in predicting the number of new students who are accepted, so it will be easier to make decisions in determining the next steps and estimating the financial matters. The prediction of the number of new student acquisition will facilitates STMIK BANJARBARU in determining the number of classes, setting schedules and others.

II. THEORITICAL BASIS

A. Prediction or forecasting

Prediction or forecasting is an important tool in an effective and efficient plan, especially in the economic field. In modern organizations, knowing the coming state is very important to look at the good or bad and aimed to prepare for the next activities (Rambe,2012).

According Heizer and Render (2009: 162), forecasting is the art and science to predict future events. This can be done by involving the retrieval of historical data and projected into the future with a form of mathematical models or predictions are subjective intuition, or using a combination of mathematical models that are tailored to the good judgment of a manager.

According Prasetya and Lukiastuti (2009: 43), forecasting is an attempt to predict the future state through state testing in the past. Forecasting relates to attempt to predict what happens in the future, based on the scientific method (science and technology) and carried out mathematically. However, forecasting activities are not solely based on scientific procedures or organized, because there is activity forecasting that uses intuition (feeling) or through informal discussions in a group.

According to Yamit, Forecasting is a prediction, projection or estimate of the level of an uncertain event in the future (Rambe 2012). According Makridakis, Forecasting is predictive values of a variable based on the known value of the variable or variables related (Rambe 2012).

According Pangestu Subagyo (1991: 1), forecasting is an activity / business to know (event) will happen in the future

regarding a particular object by using experience / historical data. According to T Hani Handoko (1994; 260) Forecasting is an attempt to predict the future state through testing in the past.

From some of these explanations, it can be concluded that forecasting is a process or method of predicting an event that will occur in the future by basing it self on certain variables.

B. Least Square method

Least Square is a method for determining the approach polynomial function y = f (x) closest to the data (x1, y1) to (xn, yn). (Basuki 2014)

In the Collins English Dictionary says that the Least Square method is the best method to determine the value of an unknown quantity related to one or more sets of observations or measurements. (Harper Collins, 1991, 1994, 1998, 2000, 2003).

According to Dr. Setijo in his module entitled "Linear Regression with Least Squares Method" said that Least Square method is an approach method which is widely used for:

1) *Regression modeling based on the equation of the discrete data points*
2) *Analysis of measurement error (model validation) (Kristalina, 2015).*

This method is most often used to predict (Y), because the calculation is more accurate. The equation of the trend line to be searched is

$$Y = a_0 + bx \quad a = (\Sigma Y) / n \quad (1)$$
$$b = (\Sigma XY) / \Sigma x^2 \quad (2)$$

with:

Y $\quad$ = periodic data (time series) = estimated trend value.

a_0 $\quad$ = trend value in the base year.
b $\quad$ = the annual average growth of value trend.
x $\quad$ = time variable (days, weeks, months or years).

To perform a calculation, it will require a specific value on the time variable (x), so that the amount of the time variable value is zero or $\Sigma x = 0$.

In this case,it will be devoted to discuss the analysis of time series with Least Square method which is divided into two cases, namely the even data case and the odd data case.

For odd n, then:

1) *The distance between the two time is rated one unit.*
2) *Above 0 is marked negative*
3) *Under 0is marked positive.*

For even n, then:
1) *The distance between the two time is rated two units.*
2) *Above 0 is marked negative*
3) *Under 0 is marked positive.*

In the data processing of odd data, the registration data of new student candidates for the past five years is required,

which are the registration data in 2010 until the registration data in 2014. In the data processing of even data, the registration data of new student candidates for the past five years is required, which are the registration data in 2009 until the registration data in 2014.

In general, linear line equation of time series analysis is:

$$Y = a + b X. \quad (3)$$
Information:
Y is the variable whichits trendis sought

X is the variable of time (years).

While, to find constant value (a) and parameter (b) is:

$$a = \Sigma Y / N \quad (4)$$
and
$$b = \Sigma XY / \Sigma X2 \quad (5)$$

III. SYSTEM ANALYSIS AND DESIGN

A. Literature Review

The previous study, conducted by Muhammad Ihsan Fauzi Rambe in 2012, examined the prediction of medicine supply using least square method which took the case study at Mutiara Hati Pharmacy, Medan. The study found that Least Square Method can be used to predict the medicine sales in the coming period based on the sales data in the previous year. Further, the analysis application can yield predictions and has minimized the forecast errors of the level of medicine sale in Pharmacies.

B. Data Requirements

The data required in the study is the data of Students accepted in STMIK BANJARBARU in 2009 to 2014.

TABLE I. DATA OF NEW STUDENTS OF STMIK BANJARBARU

No.	Year	Total
1	2009	407
2	2010	513
3	2011	528
4	2012	374
5	2013	375
6	2014	385

(Source: PMB STMIK BANJARBARU)

IV. RESULTS AND DISCUSSION

A. Odd Data Case

Before calculating the prediction of new student acquisition in 2015, some trialswere conducted in calculating the number of new student acquisition in 2012, 2013 and 2014 to determine the validity of the Least Square Method formula.

In calculating the prediction result of the number of students in 2012, the researcher used the student data in 2007 to 2011.

TABLE II.   DATA OF NEW STUDENTS IN 2007 TO 2011

No.	Year	Total
1	2007	350
2	2008	512
3	2009	407
4	2010	513
5	2011	528

(Source: PMB STMIK BANJARBARU)

The next step is determining the values of variable X, XY and X^2.

TABLE III.   DATA OF NEW STUDENTS IN 2007 TO 2011

No.	Year	Amount(Y)	X	XY	X^2
1	2007	350	-2	-700	4
2	2008	512	-1	-512	1
3	2009	407	0	0	0
4	2010	513	1	513	1
5	2011	528	2	1056	4
Total		2310		357	10

(Source: PMB STMIK BANJARBARU)

It is known that:

$\Sigma Y = 2310$

$N = 5$

$\Sigma XY = 357$

$\Sigma X2 = 10$

Then, to find the value of a:

a $= \Sigma Y/N$

a $= 2310/5$

a $= 462$

And to find the value of b:

b $= \Sigma XY/\Sigma X2$

b $= 357/10$

b $= 35.7$

After the values of a and b are obtained, then the linear line equation is:

Y$= a+bX$

Y$= 462+(35.7) X$

With the calculated equation of linear line, the number of new student in 2012 can be predicted:

Y$=462+(35.7)$ X(For the year of 2012, the value of X is 3)

so that:

Y$=462+(35.7 \times 3)$

Y$=462+107.1$

Y$=569.1$ (6)

It means that the number of new student candidates who registered in 2012 was 569 people.

The next was calculating the result of the number of new student acquisition in 2013. The data used was the data of new student in 2008 to 2012.

TABLE IV.   NEW STUDENT DATA IN 2008 TO 2012

No.	Year	Total
1	2008	512
2	2009	407
3	2010	513
4	2011	528
5	2012	374

(Source:PMB STMIK BANJARBARU)

Next was determining the values of X, XY and X^2.

TABLE V.   NEW STUDENT DATA IN 2008 TO 2012

No.	Year	Amount	X	XY	X^2
1	2008	512	-2	-1024	4
2	2009	407	-1	-407	1
3	2010	513	0	0	0
4	2011	528	1	528	1
5	2012	374	2	748	4
Total		2334		-155	10

(Source: PMB STMIK BANJARBARU)

It is known that:

$\Sigma Y \ = 2334$

$N \ \ = 5$

$\Sigma XY = -155$

$\Sigma X^2 \ = 10$

Then, to find the value of a:

a $= \Sigma Y / N$

a$= 2334/5$

a$= 466.8$

And to find the value of b:

b$= \Sigma XY/ \Sigma X^2$

b$= -155/10$

b$= -15.5$

After the values of a and b are obtained, then the equation of linear line is:

Y $= a + b X$

Y $= 466.8 + (-15.5)X$

With the calculated linear line, it can be predicted that the number of new students in 2013 is:

Y $= 466.8 + (-15.5)$ X (For the year of 2013, the value of X is 3)

Thus:

Y = 466.8– (15.5 x 3)

Y = 466.8 – 46.5

Y = 420   (7)

It means the number of candidates who registered in 2013was420.

Next was calculating the number of new student acquisition in 2014. The data used was the student data in 2009 to 2013.

TABLE VI.   DATA OF NEW STUDENT IN 2009 TO 2013

Year	Amount
2009	407
2010	513
2011	528
2012	374
2013	375
Total	2197

(Source:PMB STMIK BANJARBARU)

The next step is determining the variable values of X, XY and X^2.

TABLE VII.   DATA OF NEW STUDENTS IN 2009 TO 2013

Year	Amount	X	XY	X^2
2009	407	-2	-814	4
2010	513	-1	-513	1
2011	528	0	0	0
2012	374	1	374	1
2013	375	2	750	4
Total	2197		-203	10

(Source: PMB STMIK BANJARBARU)

It is known that:

ΣY = 2197

N   = 5

ΣXY = -203

ΣX^2 = 10

Then, to find the value of a:

a = ΣY / N

a= 2197/5

a= 439.4

And to find the value of b:

b= ΣXY/ ΣX^2

b= -203/10

b= -20.3

After the values of a and b are obtained, then the equation of linear line is:

Y = a + b X

Y = 439.4 + (-20.3)

With the calculated linear line, it can be predicted that the number of new students in 2014 is:

Y = 439.4+ (-20.3) X (For the year of 2014, the value of X is 3)

Thus:

Y = 439.4 – (20.3 x 3)

Y = 439.4 – 60.9

Y = 378.5   (8)

It means that the number of new student candidates who registered in 2014 was 378.

From the results of calculations in predicting the acquisition of the number of new students (6) (7) (8), it was found that in 2012 there was 569, in 2013 there was 420 and in 2014 there was 378. It is determined that if the deviation between the fact and the calculation with Least Square Method is >50 people, then the result is invalid. Compared to the tangible result obtained in 2012, the deviation is 34.2% (195 people), meaning that the result is invalid. In 2013, the deviation is 10% (45 people), meaning that the result is valid. In 2014, the deviation is 2.07% (8 people), meaning that the result is valid. From the three comparisons, it is found that two results are valid and one result is invalid. This means that the formula of Least Square Method is valid or closer to the truth.

Next, the calculation would be performed to predict the number of new student acquisition in 2015. In the data processing of Odd Data case, the data of new students needed is the data from the last 5 years, from 2010 to 2014.

TABLE VIII.   DATA OF NEW STUDENT REGISTRATION IN 2010 TO 2014

Year	Amount
2010	513
2011	528
2012	374
2013	375
2014	386
Total	2176

(Source: PMB STMIK BANJARBARU)

Then, determining the variable values of X, XY dan X^2.

TABLE IX.   DATA OF NEW STUDENT REGISTRATION IN 2010 TO 2014

Year	Amount(Y)	X	XY	X^2
2010	513	-2	-1026	4
2011	528	-1	-528	1
2012	374	0	0	0
2013	375	1	375	1
2014	386	2	772	4
Total	2176		-407	10

(Source: PMB STMIK BANJARBARU)

Based on Table 3, the values of a and b will be discovered. To find the values of a and b:

It is known, that:

ΣY = 2176

N   = 5

ΣXY = -407

ΣX² = 10

Then, to find the value of a:

a = ΣY / N

a= 2176/5

a= 435,2

And to find the value of b:

b= ΣXY/ ΣX²

b= -407/10

b= -40.7

After the values of a and b are obtained, then to find the equation of linear line:

Y = a + b X

Y = 435.2 + (-40.7)

With the calculated equation of linear line, the number of new student in 2015 can be calculated:

Y = 435.2 + (-40.7) X (For the year of 2015, the value of X is 3)

So that:

Y = 435.2 – (40.7 x 3)

Y = 435.2 – 122.1

Y = 313.1   (9)

It means that the numbers of new student candidates who register are 313 people.

B. Even Data Case

TABLE X.   DATA OF NEW STUDENT REGISTRATION IN 2009 TO 2014

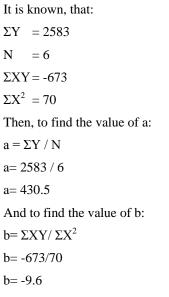| Year	Amount
2009	407
2010	513
2011	528
2012	374
2013	375
2014	386
Total	2583

(Source: PMB STMIK BANJARBARU)

The next step is determining the variable values of X, XY and X².

TABLE XI.   DATA OF NEW STUDENT REGISTRATION IN 2009 TO 2014

Year	Amount (Y)	X	XY	X²
2009	407	-5	-2035	25
2010	513	-3	-1539	9
2011	528	-1	-528	1
2012	374	1	374	1
2013	375	3	1125	9
2014	386	5	1930	25
Total	2583		-673	70

(Source: PMB STMIK BANJARBARU)

Based on Table 4, the values of a and b will be discovered. To find the values of a and b:

It is known, that:

ΣY   = 2583

N   = 6

ΣXY = -673

ΣX² = 70

Then, to find the value of a:

a = ΣY / N

a= 2583 / 6

a= 430.5

And to find the value of b:

b= ΣXY/ ΣX²

b= -673/70

b= -9.6

After the values of a and b are obtained, then to find the equation of linear line:

Y = a + b X

Y = 430.5 + (-9.6)

With the calculated equation of linear line, the number of new student in 2015 can be calculated:

Y = 430.5 + (-9.6) X (For the year of 2015, the value of X is 7)

Thus:

Y = 430.5 – (9.6 x 7)

Y = 430.5 – 67.2

Y = 363   (10)

It means that the numbers of new student candidates who register are 363 people.

From the calculation using Least Square formula (9) (10), the results showed that the prediction of the number of new student acquisition in 2015 for Odd Data is 313 people and for Even Data is 363 people. But the calculation result of the numbers of new student prediction can be damaged or fell due to several reasons, for example because of changes in government regulations, regression of high school graduates,or other reasons.

V.   CONCLUSIONS

Based on the results of the study, it can be concluded that:

1) Prediction or forecasting analysis using Least Square method can be used to predict the number of new students acquisition for the coming period based on the data of the

previous years, because the results are valid or closer to the truth.

2) From the results of calculations in predicting the acquisition of the number of new students, it was found that in 2012 there was 569, in 2013 there was 420 and in 2014 there was 378. It is determined that if the deviation between the fact and the calculation with Least Square Method is >50 people, then the result is invalid. Compared to the tangible result obtained in 2012, the deviation is 34.2% (195 people), meaning that the result is invalid. In 2013, the deviation is 10% (45 people), meaning that the result is valid. In 2014, the deviation is 2.07% (8 people), meaning that the result is valid. From the 3 comparisons, it is found that 2 results are valid and 1 result is invalid. This means that the formula of Least Square Method is valid or closer to the truth.

REFERENCES

[1] Basuki, A. (2014). Metode Least Square. taken from i http://basuki.lecturer.pens.ac.id/lecture/numerik5.pdf

[2] Beny Mulyandi, Y. I. (2010). Sales Forecasting Analysis of Fuel type Premium at the pump Heroes Bandung Asri. National Conference: Design And Application Of Technology .

[3] Cahyo Adi Basuki, I. A. (2008). Analysis of Fuel Consumption In Steam Power Plant.

[4] Geer, S. A. (2005). Least Squares Estimation. Encyclopedia of Statistics in Behavioral Science, Volume 2, pp. 1041–1045 .

[5] Handoko, T. H. (1994). Dasar – Dasar Manajemen Produksi dan Operasi. Yogyakarta: BPFE.

[6] HarperCollins. (1991, 1994, 1998, 2000, 2003). Dictionary / Thesaurus. Taken from The Free Dictionary By Farlex: http://www.thefreedictionary.com/Least-squares+method

[7] Heizer, J. d. (2009). Manajemen Operasi, Edisi 9.

[8] Kosasih, S. (2009). Manajemen Operasi - Bagian Pertama. Edisi 1. 74.

[9] Kristalina, P. (2015). Metode Least Square.

[10] Mia Savira, Nadya N.K. Moeliono, S.SOS, MBA. (2014). Sales Forecasting Analysis of generic drugs bearing (OGB) At PT. Indonesia Farma.

[11] Rambe, M. I. (2012). Perancangan Aplikasi Peramalan Persediaan Obat Obatan Menggunakan Metode Least Square. Pelita Informatika Budi Darma, Volume : VI, Nomor: 1, Maret 2014

[12] Subagyo, P. (1999). Forecasting (Konsep dan Aplikasi). Yogyakarta.

[13] Sahara, Afni. (2013). Sistem Peramalan Persediaan Unit Mobil Mitsubishi Pada PT. Sardana Indah Berlian Motor Dengan Menggunakan Metode xponential Smoothing. Informasi dan Teknologi Ilmiah(INTI), Volume : I, Nomor : 1, Oktober 2013

[14] Tanojo, E. (2007.). DeriVation Of Moving Least Squares Approximation Shape Functions And ITS Derivatives Using The Exponential Weight Function. Civil Engineering Dimension, Vol 9, No 1, 19-24, March 2007 .

[15] Widodo, J. (2008). Ramalan Penjualan Sepeda Motor Honda pada pada CV. Mitra Roda Lestari. (xii + 32 + Lampiran).