



ISSN: 2350-0328

**International Journal of Advanced Research in Science,
Engineering and Technology**

Vol. 8, Issue 12 , December 2021

Using The Least Square Method for Estimation Number of New Students

Rina Candra Noor Santi, Sri Eniyati, Veronica Lusiana

Department of Computer Science/Informatics, Information Technology and Industry Faculty, Stikubank University,
Indonesia.

ABSTRACT: The number of new students admitted to several universities, both public and private, will affect the teaching and learning process at these universities. Knowing the number of new students is one thing that can be used for planning materials in the teaching and learning process, because it will be related to the number of existing lecturer ratios.

This study aims to predict the number of new students for the next year using the Least Square method. The data used is the number of new students in the previous 6 years.

The results of this forecasting function to help get the results of forecasting the number of new students in the future, so that it can assist in planning the teaching and learning process

KEYWORDS: interpolation, least square, forecasting, estimation

I. INTRODUCTION

Forecasting is an activity of implementing a model that has been developed in the future [1]. Forecasting is the process of estimating future demand related to aspects of quantity, quality, time of occurrence, and the location that requires the product or service concerned.

New student is a status that is carried by students in their first year of college. Entering the world of college is a big change in one's life. Usually, individuals experience many changes in their first year of college when entering college, either in public or private universities.

The number of new students accepted in several universities, both public and private, will affect the learning and teaching process at these universities. Knowing the number of new students is one thing that can be used for planning materials in the teaching and learning process, because it will relate to the number of existing lecturer ratios.

Therefore, it is better to predict the number of new students to prepare everything. Prediction or forecasting in the modern era as it is today is widely used, because by using existing forecasts and predictions, it will help to know the number of new students who will come.

Many forecasting methods have been used, including WMA (Weighted Moving Average), Single Exponential Average, ARIMA, Linear Interpolation, Quadratic Interpolation, Lagrange Interpolation, Newton's Interpolation and others. All of these methods have their respective advantages and disadvantages.

With these several forecasting methods, the author tries to predict the number of students with one method, namely by using the Least Square Method. Many say that using this method is easier to understand and simpler.

The aim of this research is to predict and predict the number of new students for the coming year using the Least Square method. And the benefit of this research is to prepare and plan the teaching and learning process to be better by looking at the number of lecturer ratios.

**II. LITERATURE REVIEW**

There are several forecasting methods that can be used based on their nature:

a. Qualitative Forecasting

Qualitative forecasting is forecasting based on the opinion of a party and the data cannot be represented explicitly into a number or value. The results of forecasts made are very dependent on the person who compiled them. This is important because the forecasting results are determined based on intuitive thinking, opinions and knowledge and experience of the compilers.

b. Quantitative Forecasting (Statistical method)

Quantitative forecasting is forecasting based on past quantitative data and can be made in the form of numbers.

Quantitative forecasting can only be used if there are three conditions as follows:

1. Information about past circumstances.
2. The information can be quantified in the form of numerical data.
3. It can be assumed that some aspects of the past pattern will continue in the future.

There are several forecasting models that are classified as quantitative methods, namely:

a. Time series models

The Time Series method deals with the values of a variable that are set periodically over the time period in which the demand forecast is projected. For example weekly, monthly, quarterly, and yearly, depending on the wishes of the parties who forecast this demand. This method is solely based on data and past conditions. If conditions in the future are stable enough in the sense that there are not many significant changes with past conditions, this method can provide fairly accurate forecasting results.

b. Linear Trend Model

The Linear Trend has an equation which in general can be expressed as follows:

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

Description :

Y : trend value in a certain period

X : time period

a : the intercept of the trend equation

b : the slope or gradient coefficient of the trend equation which shows the magnitude of a unit change in X

There are four methods that can be used to construct or determine a linear trend, namely:

1. Free Method (Freehand Method)
2. Semi Average Method (Semi Average Method)
3. The Moving Average Method
4. Least Square Method

Least Square Method

The least square method or commonly called the least squares method was discovered by Carl F. Gauss (a famous German mathematician and physicist, 17th century) when he was 18 years old, and his work is still used today as the best method. to determine the linear relationship of two data variables.

Least square is the method used to determine the trend equation of the data because this method produces data mathematically. In this case, it will be more devoted to discussing the analysis of the least square method which is divided into two cases, namely the case of even data and odd data.

The principle of the least squares method is to minimize the sum of the squared deviations (difference) in the value of the independent variable (Y_i) with the trend / forecast value (Y') or $(Y_i - Y')^2$ being minimized.



With the help of calculus, namely the partial deviation, $(Y_i - Y')^2$ is minimized, then two normal equations will be obtained as follows (Joko Widodo, 2008):

$$\sum Y_i = n.a + b.\sum X_i \dots\dots\dots(2)$$

$$\sum X_i Y_i = a.\sum X_i + b.\sum X_i^2 \dots\dots\dots(3)$$

By solving these two normal equations simultaneously, the values of a and b of the trend equation $= a + b$ can be calculated.

In order to make the calculation simpler, the code for the value of X (years) is attempted in such a way that $= 0$, so that the normal equation above can be simplified as follows (Joko Widodo, 2008):

$$a = \sum Y_i / n \dots\dots\dots(4)$$

$$b = \frac{\sum X_i Y_i}{\sum X_i^2} \dots\dots\dots(5)$$

After the values of a and b are calculated using the above formula, the equation for the linear trend value can be arranged as follows (Joko Widodo, 2008):

$$Y = a + bX \dots\dots\dots(6)$$

Y = trend value in a certain period

a = intercept, i.e. the magnitude of the value of Y if the value of X = 0

b = the slope of the trend line, i.e. the change in the Y variable for every change in one unit of the variable X

X = time period

To make the value of $= 0$ depending on the number of year data, namely even and odd, the guidelines are as follows: (Budiasih Yanti, 2012)

If the number of data years is not divisible by two, which is odd, then a scale of $x = 1$ year is used. Then the base year is placed in the middle year.

The purpose of forecasting is to forecast demand and independent demand items in the future, to get a forecast that can minimize forecasting errors (Forecast Error) which can be measured by Mean Absolute Error (MAE) and Mean Squared Error [3]. With this sales forecast, it means that the company's management has got a picture of the company in the future, so that the company's management will get very meaningful input in determining company policies [1].

Several studies have been used as literature review material, including in his research, namely predicting stock prices so that they are free from risks such as losses.

Therefore, it is necessary to analyze stock price estimates to reduce this risk by using Lagrange interpolation and Newton's interpolation and then compare the two interpolations.

The application of Newton's interpolation method is also used to predict the number of HIV sufferers. And from the results of his research the number of HIV sufferers from 2014-2016 has increased [4].

III. METHODOLOGY

In this study, the type of data used is quantitative data. Sources of data used in this study is primary data. Primary data sources are data obtained directly on the object of research.

To obtain and collect the data needed, the author uses several methods as follows:

1. Observation

Observation, namely direct observation and recording of data and information in the public relations department of Semarang.

2. Documentation

Documentation is looking for documents that have to do with discussing the problem to complete the data needed in writing

3. The data obtained from the two data collection methods above are primary data.

4. Literature Review

Literature review, namely data obtained from theoretical knowledge obtained in full, both from lecture materials, relevant reference books and from browsing results on the internet related to this research.

The stages / flow in the research to be carried out can be seen from the following picture:

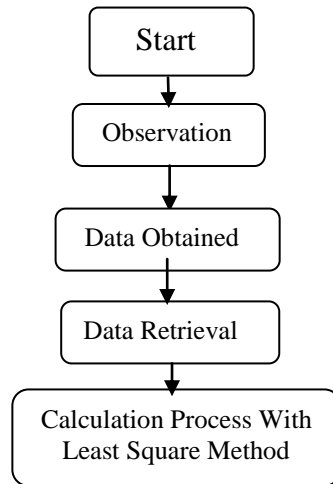


Fig 1. Research Flow

This research was conducted using data from the private university in Indonesia, using observation methods and data feasibility studies. The method used begins with analyzing the data obtained.

IV. RESULTS AND DISCUSSION**a. Application of the Least Square Method**

Trying to predict the results of the optimal number of students, using the least square method.

The formula is as follows:

$$Y = a + bx \quad \dots\dots\dots (7)$$

Description :

Y : Number of new students

a and b : Coefficient

x / t : specific time

In determining the value of x / t, alternative techniques are often used by giving a score or code.

In this case, the data is divided into two groups, namely:

a. The data is even, then the t-value score is: ..., -5, -3, -1, 1, 3, 5, ...

b. Odd data, then the t-value score is: ..., -3, -2, -1, 0, 1, 2, 3, ...

Then to find out the coefficients a and b are searched with the equation

$$a = \frac{\sum Y}{n} \quad \dots\dots\dots (8)$$

$$b = \frac{\sum t Y}{\sum t^2} \quad \dots\dots\dots (9)$$



The new student table can be seen as follows:

Table 1. Number of New Students 2015-2019

x	2015	2016	2017	2018	2019
f(x)/y	338	354	314	276	256

Description:

x = is the independent variable (th)

f(x) or y = is the number of new students

b. The calculation results

In this case, we try to calculate the estimated number of new students for the next 5 years from the existing student table.

The results of calculations using Microsoft Excel are as follows:

No	Years	Value (Y)	X	X ²	XY	Y'
1	2015	338	-2	4	-676	348
2	2016	354	-1	1	-354	331,8
3	2017	314	0	0	0	315
4	2018	276	1	1	276	299,4
5	2019	256	2	4	512	283,2
	Total	1538	0	10	-242	

So from the results above, we can find the values of a and b. With the following rules:

a = amount of data y : lots of data

b = sum of xy: sum of x²

From the rules above, the value of a = 1538/5 = 307.6

b = -242/10 = -24.2

So with Y = a+bx then Y = 307.6 + (-24,2) x

From the known equation, namely Y = 307.6 + (-24.2) x , then the forecast for the next five years can be known.

Forecasting results for the next five years are as follows:

Year	Value
2020	243
2021	218,8
2022	194,6
2023	170,4
2024	146,2



ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 8, Issue 12 , December 2021

From the results above, the error value can be found.

The meaning of error or commonly called error in the numerical method is the difference between the resulting value and the value generated by the numerical method. In the numerical method, the result obtained is not exactly the same as the true value.

There will always be a difference, because the results obtained with the numerical method are the results obtained by iterating (looping) to approach the actual value. However, it does not mean that the results obtained by the numerical method are wrong, because the error can be reduced as small as possible so that the results obtained are very close to the true value or it can be said that the error is close to zero.

The way to get the error value is as follows:

$$\text{Error 2020} = |\text{true value} - \text{approximate value}| / \text{true value} \\ = |296 - 243| / 296 = 0.17$$

$$\text{Error 2021} = |\text{true value} - \text{approximate value}| / \text{true value} \\ = |146 - 218| / 146 = 0.47$$

So the error value in 2020 still tends to be small. But the error in 2021 tends to be large.

V. CONCLUSION

The conclusions obtained from this research are:

1. Forecasting systems can assist in forecasting to determine the number of students in the future.
2. From the results of the research above, it is found that the application of the least square method may be used to predict the number of new students, if the increase and decrease in students is not so drastic. However, for the data taken here, it appears that the increase and decrease in each year is not linear, so that the application to determine or predict the number of new students using the least squares method is not valid.

REFERENCES

- [1] Aritonang, Lerbin R., "Business Forecasting", Ghalia Indonesia Publisher, 2013
- [2] Ahmad Chamami, "Interpolation to predict the number of FKIP PGRI students", Kediri, 2016
- [3] Gaspersz, "Production Planning and Inventory Control", Jakarta, Gramedia Indonesia, 2015
- [4] Munir R., "Numerical Method", 2nd Edition, Bandung, Bandung Informatics, 2003
- [5] Tony Yulianto, Nur Ita U, Rica Amalia, "Forecasting HIV Using Lagrange and Newton Interpolation", Zeta math Journal, Vol. 2 No. 1 May 2016
- [6] Sri Eniyati, "Use of the Lagrange Method for Estimating the Number of New Students", Proceedings, SENDIU, June 2019
- [7] Saefurrohman, "Forecasting new students using the Newton Raphson Method", Unisbank Internal Research, 2020.