Application of Data Mining with the Least Square Method to Predict Web-Based Drug Inventory

Abdul Halim*, Safwandi, Fajriana

Department of Informatics, Faculty of Engineering, Universitas Malikussaleh, Aceh, Indonesia

*Corresponding author E-mail: abdul.180170090@mhs.unimal.ac.id

Abstract

Drug supplies are an important aspect because of their large value and large quantity and are an important factor in supporting health services in community health centers. Ineffective drug management, especially in terms of needs planning, can lead to excess or shortage of stock. Both conditions have negative impacts, such as budget waste, drug expiration, or even disruption of patient services due to unavailability of drugs. At the Pante Bidari Health Center UPTD, the drug needs planning process is still carried out manually or based on rough estimates without using sophisticated technology. This study aims to design and build a web-based drug inventory prediction system using the Least Square method. The Least Square method was chosen because it is able to carry out the forecasting process quickly and with good results. In this study, the type of data obtained is drug usage data, data is grouped based on each supplier, from the health center information system during a certain period. After going through the pre-processing and calculation stages, the predicted values are calculated and displayed through a web-based system designed to be easy to use by health center officers. The web system developed in this study uses PHP as the programming language and MySQL as the database, implementing the Least Square method effectively. The results of this study are a drug usage prediction application for the future, applying the Least Square method, which displays drug usage data over a certain period. The system will present the data in the form of a table. Based on testing the drug usage data for Acyclovir Cream 5 mg from January 2023 to August 2024, the prediction result for the following month, September 2024, is estimated to be 38.415, which is rounded to 38 units of the drug.

Keywords: *Data Mining, Least Square, Health Center, Drug Supplies, Web-Based.*

1. Introduction

Puskesmas (Community Health Center) is a health service facility that supports community health initiatives and individual health, emphasizing preventive and promotive initiatives to achieve ever-increasing public health goals in community work. Health services can be productive if they are in line with community needs, one of which is the availability of medicines [1].

UPTD Pante Bidari Health Center is a public health agency located in Meunasah Tunong Village, Pante Bidari District, East Aceh Regency. This institution contains all information and data related to general health, health development, and development performance, such as health data, health facilities, health workers, and specific health program data. As part of public health in Pante Bidari District, health services must not only be provided to the community as a whole, However, the health service delivery system must also be examined and evaluated, both in real time and during the drug transaction process.

Drug supplies are one of the factors that must be considered and managed by the health center. Drug supplies are an important aspect because of their large value and large quantities and are an important component in the patient healing process.

Several studies using the Least Square method on blood stock inventory prediction. Factors or parameters that influence the demand for blood products are the environment, disease symptoms, and differences in the immune system of each blood type. In addition, different types of blood products have varying amounts of demand, depending on the level of blood need for health. With a MySQL database and a system designed using the PHP programming language, the Cirebon City PMI Blood Transfusion Unit can use it to make it easier for users to handle data and blood and blood stock report data for each period [2].

Another study on LED lamp sales predictions using the Least Square method. The predicted sales results for LED downlights in July 2021 were 13 units, indicating that there was no increase or decrease in the number of sales compared to the previous year. The Least Square method is suitable for predicting sales using time series data, and the MAPE calculation has an error rate of around 8.0744%. As a result, it can be concluded that the performance of the forecasting model is very good, and the applied data mining system can provide new information that is very useful for the company [3].

Other research on inventory forecasting using the Least Square method. Companies can predict how much sales will occur in the coming period, allowing them to choose how much stock of their products to keep on hand [4].

Another study also on predictions for soybean raw material requirements using the Least Square Method. This forecasting system can regulate the amount of soybean raw materials because this forecasting system has accuracy in predicting the need for soybean raw materials. The soybean raw material requirement forecasting system can be improved by using the Least Square method. The results of the calculation of the need for soybean raw materials in January 2023 were 5,376 kg, in February 2023 it was 5,433 kg, and in March 2023 it was 5,489 kg. The error rates obtained were MSE 69196, RMSE 263, MAD 214, and MPE 0.04% [5].

Based on several case study findings and several completed studies, the author conducted a study entitled "Application of Data Mining with the Least Square Method to Predict Web-Based Drug Inventory".

2. Literature Review

2.1. Data Mining

Data mining is an academic field used to solve problems involving the analysis of large databases by combining statistical techniques, machine learning, data visualization, pattern analysis, and databases [6]. Data mining is the process of extracting useful information and patterns from large amounts of data. The data mining process consists of data collection, data analysis, data extraction, and data statistics. The term data mining is also often referred to as knowledge discovery [7].

Data Mining is a procedure for extracting useful information from large databases that must be carefully examined to produce new information that can help in developing a decision [8]. Data mining is the process of transforming data sets into information with previously unknown implicit (non-obvious) potential [9]. The goal of data mining is to analyze data using certain methods and then transform the data into a structure that can be understood for further use [6].

Data mining is effective for use not only in business environments but also in other environments such as weather forecasting, medical science, health care, transportation, insurance, and government [10].

2.2. Drug Supplies

Drug inventory is the stock of drugs stored in health facilities to meet patient needs. Good management of drug inventory is essential to prevent shortages and ensure timely availability of drugs [11]. Good supplies can improve the quality of health care and reduce the risk of missing opportunities for effective treatment.

2.3. Predictions

The definition of prediction is the same as forecast or estimate. Forecasting is a technique for analyzing future events that is carried out using quantitative and qualitative methods to estimate future events using historical data to minimize negative impacts. This method is applied to the production process so that business owners will be more cooperative in the production process. This is because this method can provide the best output, Therefore, it is hoped that the risk of problems arising in the production process can be reduced as low as possible [12].

2.4. Least Squares Method

One of the analysis methods used to see the trend of time series data is called Least Square [13]. The Least Square method is an algorithm in the form of time series data that uses data from the previous period which is used to predict sales in the next period or the coming month. The least squares method is used to forecast the next period's performance using previous time data [14].

The Least Square method is divided into two categories, namely odd data and even data. Analysis using the Least Square method can be summarized as follows:

$$Y' = a + b \cdot X \tag{1}$$

Information:
a = Trend value in the base period
b = Average growth of trend value for each period
Y' = Prediction result value
X = Time variable

The steps for implementing the Least Square method will be explained as follows [15]:
1. Data Sample (Variable Y)
   The author obtained data from the UPTD Pante Bidari Health Center. The type of data obtained is drug use data (Y).
2. Determining the Time Variable (Variable X)
   In determining the value of X, alternative techniques are often used by providing codes or scores. Therefore, the number of time variables is zero or $\sum X = 0$. In this process, the data is analyzed into two groups, namely:
   a. Even data or period data (n) is an even number
       So the value of X in the middle data (n) starts from -1 and 1, the distance between two times is given a value of two units. Above is given a negative sign, and below is given a positive sign. So the value of X is: …,-5, -3, -1, 1, 3, 5, ...
   b. Odd data or data period (n) is odd in number
       So the value of X in the middle data (n) starts from 0, the distance between two times is given a value of one unit. Above is given a negative sign, and below is given a positive sign. So the value of X is: …,-3,-2, -1, 0, 1, 2, 3, ...
3. Calculating the Value of X^2 and the Value of X.Y
   The value of X is squared based on the data for each month. Calculate the value of X.Y where Y is the usage data and X is the time index, both multiplied.
4. Calculating the Value of Coefficient a and the Value of Coefficient b
   The next step is to calculate the value of Coefficient a and Coefficient b using the following equation:
   $$a = \frac{\sum Y}{n} \tag{2}$$

$$b = \frac{\sum X.Y}{\sum X^2}$$
(3)

5. Calculating Predicted Value (Y')
   The next step is if the coefficient a and coefficient b values that have been obtained are entered into Equation 1, where the X value sought for the following month is:
   Y' = $a + b . X$ (4)

   Information:
   a = Trend value in the base period
   b = Average growth of trend value for each period
   n = Number of period data
   ∑ = Cumulative amount
   Y′= Prediction result value
   Y = Drug use variables
   X = Time variable

2.5. Hypertext Preprocessor
PHP is an Open Source application designed to make it easier to manage MySQL. With PhpMyAdmin, you can create databases, create tables, insert, update, and modify data using a graphical user interface (GUI) and more easily without having to manually check SQL syntax [16].
PHP is a programming language that is widely used to create and develop websites and can be used with HTML. PHP is a language used in HTML documents which is used for server side work [17].

2.6. Database
A database is a collection of information organized in a computer so that it can be accessed using computer software to analyze the information [18].
A database is an application that stores a collection of data. Each database has the tools needed to create, retrieve, modify, search, and summarize the data included in it [19].

2.7. My Structure Queri Language
MySQL can be called a Relational Database Management System (RDBMS), the relationship between tables that contain data in a database. MySQL is an open source database management system that allows users to import, export, and manage data efficiently [20].

3. Research Methods

3.1. Place And Time of Research
This research was conducted at the UPTD Pante Bidari Health Center located in Meunasah Tunong Village, Teupin Bate Street, Pante Bidari District, East Aceh Regency. UPTD Pante Bidari Health Center is one of the health institutions owned by the regional government which serves as a public health center in Pante Bidari sub-district. This research started on October 1st until October 10th 2024.

3.2. Research Methods
1. Data Collection
   Data collection techniques used in writing reports include observing the objects being studied, interviews with related parties, and collecting previous data.
2. System Analysis
   The system analysis created is a drug inventory prediction system based on the amount of drug use per month using the Least Square method.
3. System Design
   System design is carried out starting from the design of the overall system architecture, which includes database design, system menu structure design, application workflow design, and output design.
4. System Implementation
   System implementation using PHP programming language and MySQL database. Implementation in the form of an application interface display to make it easier for users to use the application being built.
5. System Testing
   System testing is performed to identify potential errors or bugs in the system..
6. Evaluation and Conlusion
   Evaluate whether the software has met expectations or not. The conclusions drawn must be logical and consistent with the data collected and the results of the tests carried out.

3.3. System Schematic
The following is a diagram of a drug inventory prediction system using the Least Square method. The system schematic steps can be seen in the image below:
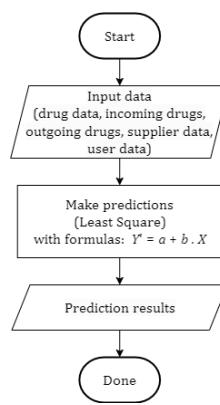
Fig 1. System Schematic

Information:
1. Start the system.
2. Inputting data, starting from drug data, incoming drugs, outgoing drugs, supplier data, and user data.
3. Carry out the calculation process using the Least Square method, with the formula: $Y' = a + b . X$.
4. Displays prediction results.
5. System complete.

4. Results and Discussion

The results of the study are in the form of a web-based application that can predict future drug supplies based on drug usage by implementing the Least Square method at the Pante Bidari Health Center UPTD. This method will produce a prediction of the amount of drug use each month at each supplier at the Pante Bidari Health Center UPTD.

4.1. System Analysis

The system created is a drug inventory prediction system based on the amount of drug use per month using the Least Square method. Analysis is performed after the system planning stage and before the system design stage. This system will require input in the form of sample data from the amount of drug usage in a certain period grouped by supplier. The sample will be processed using the Least Square method which will produce a prediction of usage for the next month. The output results, the system will display data in the form of a table that describes the movement of drug usage and movement of drug predictions.

4.2. Manual Calculation of Least Square Method

The data sample used is a sample of data on the use of National Health Insurance (JKN) drugs with the type of drug Acyclovir Cream 5 mg from January 2023 to August 2024.

To determine the value of X, look at the number of period values (n) used in the prediction. The number of period values (n) used in this study is around 20 months, so the X value is even.

Calculate the value of X squared based on each month's data. Next, calculate the value of X.Y where Y is the usage data and X is the time index, both of which are multiplied. After all the values are known, then add up (Σ) all the values, for more details see the table below:

Table 1. Calculation Table

Period	Usage (Y)	X	X^2	X.Y
January 2023	64	-19	361	-1216
February 2023	37	-17	289	-629
March 2023	17	-15	225	-255
April 2023	0	-13	169	0
May 2023	0	-11	121	0
June 2023	0	-9	81	0
July 2023	0	-7	49	0
August 2023	0	-5	25	0
September 2023	0	-3	9	0
October 2023	0	-1	1	0
November 2023	0	1	1	0
December 2023	0	3	9	0
January 2024	0	5	25	0
February 2024	0	7	49	0
March 2024	5	9	81	45
April 2024	59	11	121	649
May 2024	95	13	169	1235
June 2024	43	15	225	645
July 2024	59	17	289	1003
August 2024	39	19	361	741
TOTAL	418		2660	2218

Find the trend value a and trend value b using equation 2 and equation 3.

$$a = \frac{\sum Y}{n} \quad = \frac{418}{20} \quad = 20,9$$

$$b = \frac{\sum X.Y}{\sum X^2} \quad = \frac{2218}{2660} = 0,833835$$

After obtaining the trend value a and trend value b, then to calculate the predicted drug supply for the following month, namely September 2024, the value of X is 21.

$$Y' = a + b.X$$

$$= 20,9 + 0,833835\ (X)$$

$$= 20,9 + 0,833835\ (21)$$

$$= 20,9 + 17,51053$$

$$= 38,4105$$

So, for the sample prediction of National Health Insurance (JKN) Drug usage data with the type of Acyclovir Cream 5 mg in September 2024, it is estimated to be 38.4105 or rounded up to 38 Drugs.

4.3. System Implementation

1. Login Page

   The login page is a page that displays a form for login access to the system. In it there is a place to enter a username and password, which must be filled in to access the next page.
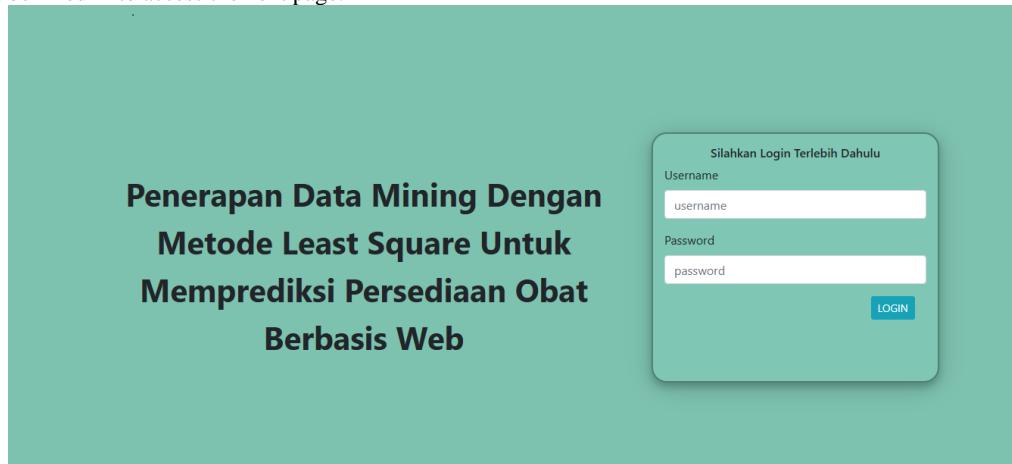

Fig 2. Login Page

2. *Dashboard Page*

   The dashboard page is the first page after the login page. This page displays a summary of the data in the database with the aim of being a brief summary of the data that has been processed.
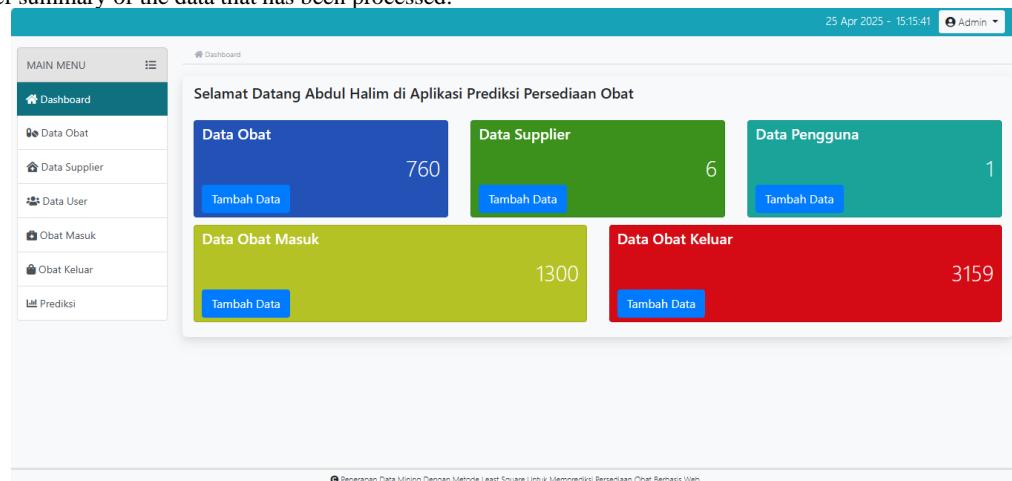

Fig 3. Dashboard Page

3. Prediction Page

   The prediction page is a page for predicting future drug supplies and the system will display the prediction results.
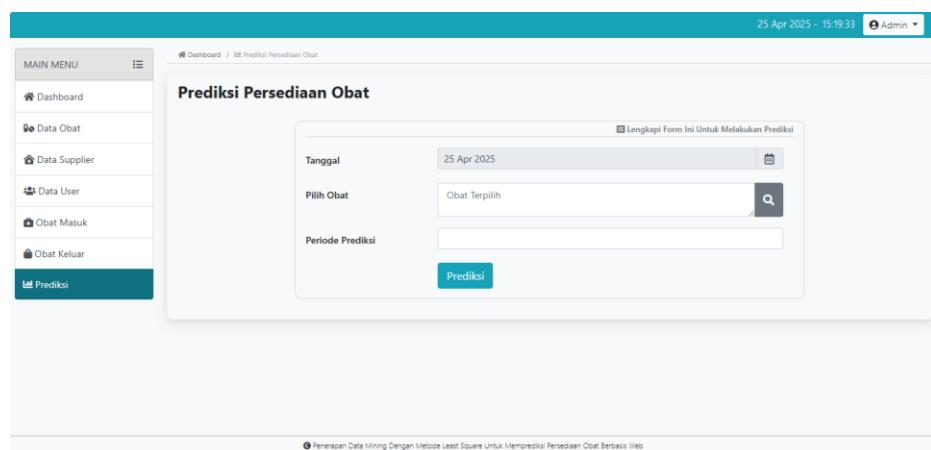
Fig 4. Prediction Page

5. Conclusion

The conclusion of the research entitled Application of Data Mining with the Least Square Method to Predict Web-Based Drug Inventory is as follows:

1. The system uses PHP as a programming language and MySQL as its database, which applies the Least Square method to the system, and can be used effectively to predict the availability of a drug.
2. This prediction system is developed using web-based technology and has an easy-to-use interface. This system makes it easy for users to enter data, view prediction results, and check reports. The use of a web-based system also offers flexibility in terms of accessibility because it can be accessed easily, thus increasing operational efficiency.
3. Based on historical data on drug use, the Least Square method can be used to determine patient needs. Through the use of historical data, the system can produce accurate predictions in estimating future needs, which can be used as a guide in drug inventory research.
4. Based on testing on the data on the use of Acyclovir Cream 5 mg from January 2023 to August 2024, the predicted results for the following month, namely September 2024, are estimated to be 38,415 or rounded up to 38 drugs.

References

[1]   Kementerian Kesehatan Republik Indonesia. (2019). *Tentang Pusat Kesehatan Masyarakat*. Jakarta: Kemenkes RI.
[2]   M. Hatta and A. F. Fitri, "Sistem Prediksi Persediaan Stok Darah Dengan Metode Least Square Pada Unit Transfusi Darah Studi Kasus PMI Kota Cirebon," Jurnal Ilmiah Ilmu Komputer, vol. 6, no. 1, pp. 19–23, 2020, doi: 10.35329/jiik.v6i1.130.
[3]   R. M. Fauzi and D. I. Mulyana, "Implementasi Data Mining Menggunakan Metode Least Square untuk Memprediksi Penjualan Lampu LED pada PT. Sumber Dinamika Solusitama," Jurnal Sosial dan Teknologi, vol. 1, no. 8, pp. 907–919, 2021, doi: 10.59188/jurnalsostech.v1i8.182.
[4]   T. Yasmi, A. Erianda, and R. Afyenni, "Implementasi Metode Least Square untuk Peramalan Persediaan Barang pada Sistem Inventori CV. Tree Jaya Perkasa," Jurnal Ilmiah Teknologi Sistem Informasi, vol. 3, no. 4, pp. 137–142, 2022, doi: 10.62527/jitsi.3.4.100.
[5]   D. N. Fazira, N. Mulyani, and E. Rahayu, "Prediksi Kebutuhan Bahan Baku Kedelai Di Pabrik Tahu Buk Iyem Menggunakan Metode Least Square," Jurnal Pendidikan Teknologi Informasi, vol. 3, no. 2, pp. 341–451, 2023, doi: 10.51454/decode.v3i2.203.
[6]   Werdiningsih, I., Nuqoba, B., & Muhammadun, S. (2020). Data Mining Menggunakan Android, Weka, dan SPSS. Airlangga University Press.
[7]   Arhami, M. & Nasir, M. (2020). Data Mining: Algoritma dan Implementasi. Penerbit Andi: Banda Aceh.
[8]   Suntoro, J. (2019). DATA MINING: Algoritma dan Implementasi dengan Pemrograman php (Edisi 1). Jakarta: PT. Elex Media Komputindo.
[9]   Muflikhah, L., Ratnawato, D. E., & Putri, R. R. M. (2021). Buku Ajar: Data Mining. UB Press: Malang.
[10] Agrawal, C. P., & Agrawal, M. (2017). Introduction to Data Mining. Educreation Publishing: New Delhi.
[11] S. Mulyani, "Manajemen Persediaan Obat di Fasilitas Kesehatan," Jurnal Manajemen Kesehatan, vol. 12, no. 1, pp. 45–52, 2020.
[12] R. A. Walangadi and I. S. Kumala, "Prediksi Penjualan Motor Dengan Menggunakan Metode Least Square," Jurnal COSPHI, vol. 3, no. 2, 2019.
[13] M. Sadli and Safwandi, "Implementasi Sistem Cerdas Least Square Dalam Meramalkan Pemenuhan Kebutuhan Stok Listrik Di Kota Lhokseumawe," Jurnal Ecotipe, vol. 4, no. 2, pp. 21–29, 2017, doi: 10.33019/ecotipe.v4i2.8.
[14] C. A. Putra and A. R. Ardhana, "Implementasi Algoritma Least Square Untuk Pembuatan Sistem Prediksi Prof-it Bulanan," SCAN, vol. 18, pp. 13–17, 2023, doi: 10.33005/scan.v18i1.3829.
[15] D. Suwardiyanto, M. N. Shodiq, D. H. Kusuma, and T. O. Sari, "Sistem Prediksi Kebutuhan Obat di Puskesmas Menggunakan Metode Least Square," Jurnal Pengembangan IT (JPIT), vol. 4, no. 1, pp. 75–80, 2019, doi: 10.30591/jpit.v4i1.1085.
[16] Madcoms. (2016). Kupas Tuntas Adobe Dreamweaver CS6 dengan Pemograman PHP dan MySQL. Yogyakarta: Andi Publisher.
[17] B. Abimanyu, R. Safe'I, and W. Hidayat, "Analisis Kerusakan Pohon Di Hutan Kota Stadion Kota Metro Provinsi Lampung," Jurnal Hutan Pulau Kecil, vol. 3, no. 1, 2019. [Online]. Available: https://ojs3.unpatti.ac.id/index.php/jhppk/article/view/1317
[18] Abdulloh, R. (2018). 7 in 1 Pemograman Web Untuk Pemula. Jakarta: PT. Elex Media Komputindo.
[19] Enterprise, J. (2017). Otodidak Pemograman Database dengan Visual Basic. Jakarta: PT. Elex Media Komputindo.
[20] Sukmawati, R. (2022). Pengantar Basis Data: Konsep dan Implementasi MySQL. Yogyakarta: Penerbit Andi.