

Data Mining

Application of Data Mining in Predicting Demand for Blood Bags Based on Blood Type Using the Single Moving Average Method (SMA)

Case Study: PMI UDD, Langkat Regency

Aqilah Tahara ^{1*}, Irvan ²

¹ Department of Information Technology, Faculty of Computer Science and Information Technology, Universitas Muhammadiyah Sumatera Utara, Medan, 20238, North Sumatra, Indonesia

² Department of Mathematics Education, Faculty of Teacher Training and Education, Universitas Muhammadiyah Sumatera Utara, Medan, 20238, North Sumatra, Indonesia

ARTICLE INFORMATION

Received: Aug 5, 2025
Revised: Oct 16, 2025
Available Online: Jan 31, 2026

KEYWORDS

Data Mining
Blood Group
Blood Demand Prediction
Single Moving Average (SMA)
UDD PMI Langkat

CORRESPONDENCE (*)

Phone: +62 821-7529-2615
E-mail: aqilah tahara459@gmail.com

A B S T R A C T

Blood stock availability at the Blood Donation Unit (UDD) PMI Langkat Regency is often disrupted due to unpredictable demand and reliance on voluntary donors, leading to risks of shortages or surpluses, especially for certain blood groups. This study applies a data mining method using the Single Moving Average (SMA) approach to predict the demand for blood bags based on blood groups (A, B, AB, and O) over the past year. The forecasting process calculates the average demand over a specific period and measures accuracy using MAPE, MSE, and MAD. The results show that the SMA method provides reasonably accurate predictions, with blood group O having the highest average demand of 161 bags per month, followed by blood groups A, B, and AB. An average MAPE value below 10% indicates that this method is effective for blood stock planning at UDD PMI Langkat, helping to optimize blood inventory management and minimize the risk of shortages or surpluses.

INTRODUCTION

Based on data from the Indonesian Red Cross (PMI), the requirements to become a donor are being physically and mentally healthy, aged 17-65 years, having a minimum body weight of 45 kg, blood pressure for systole 100-170, diastole 70-100, hemoglobin levels 12.5g%-17.0g%, and a minimum donor interval of 12 weeks or 3 months since the last blood donation [1,2,3]

The demand for blood is expected to continue to increase in line with rapid advances in medical science, particularly in organ transplant procedures. Hospitals are the places with the highest demand for blood. Blood transfusions are needed by patients experiencing blood loss due to accidents, patients undergoing major surgery, and patients with blood diseases such as leukemia, hemophilia, and thalassemia [4,5].

The problem of blood bag demand at the PMI Langkat Regency Blood Donation Unit (UDD) arises from difficulties in controlling blood supplies. Among other things, the availability of blood stock at the PMI Langkat Regency Blood Donation Unit relies on voluntary donors, which often leads to blood shortages, especially for rare blood types like AB. This results in the inability to meet the blood needs of the people of Langkat Regency and other areas. Another problem

is the presence of expired blood due to prolonged storage. If blood has expired, it cannot be used for transfusions and must be disposed of according to SOP.

Based on the aforementioned problems, the application of advanced analytical techniques, such as data mining, can provide a solution for predicting blood demand. Data mining is the process of extracting large amounts of data (Big Data) through interesting patterns that are challenging, implicit, and previously unknown, and have the potential to become useful information. The main idea of data mining is to build computer programs that automatically filter databases and search for rules or patterns using statistical techniques and mathematical models [6,7].

Forecasting is the process of estimating or predicting the quantity needed in a future period [8]. The method used in forecasting blood bag demand based on blood type is the Single Moving Average (SMA) method, which is a forecasting method carried out by collecting historical data and taking a group of observed values, finding the average value as a forecast or prediction for the next period, with a benchmark for the level of error using the Mean Absolute Deviation (MAD), Mean Square Error (MSE) and Absolute Percentage Of Error (MAPE). This method is very useful in analyzing data that has fluctuations or patterns that can be predicted over time.

In the context of this research, the Single Moving Average (SMA) method was chosen because of its ability to smooth data and identify underlying trends without requiring specific assumptions. This is particularly relevant for accurate stock predictions, as maintaining blood stock availability is crucial.

Blood demand data from the Indonesian Red Cross (PMI) Blood Unit (UDD) in Langkat Regency can be used to predict blood supply demand for the following month. Focusing on blood type, which plays a crucial role in transfusion compatibility, this study aims to analyze and predict blood bag demand at the PMI UDD in Langkat Regency. This research aims to gain a better understanding of blood demand and provide useful recommendations for increasing blood availability in the area, with the hope of saving more lives through adequate blood availability.

METHOD

Forecasting

Forecasting is the process of predicting what will happen in the future based on historical data and analysis of current trends. In the business world, forecasting is used to help companies anticipate potential needs, opportunities, and challenges, thus supporting more effective and efficient strategic decision-making.

Single Moving Average (SMA) Method

A single moving average, or simply a moving average, is a forecasting method that takes a group of observed values and finds the average value as a forecast for the future period.

The single moving average method uses a number of new actual demand data to generate a forecast value for future demand. The single moving average method has two unique characteristics [9,10] :

1. To determine the forecast for the future period, historical data from a specific period is required.
2. The longer the moving average period, the more pronounced the smoothing effect in the forecast, resulting in a smoother moving average.

RStudio

RStudio is a free and open-source integrated development environment (IDE) for the R programming language, used for statistical computing and graphics. RStudio is available in two main editions: RStudio Desktop (a local application) and RStudio Server (accessed through a browser on a remote Linux server). RStudio supports Windows, macOS, and Linux.

Research Design

In this study, a quantitative approach was used to analyze historical data on blood bag demand based on blood type (A, B, AB, and O) using the Single Moving Average (SMA) method, thus providing accurate predictions.

Quantitative approaches focus on collecting and analyzing numerical or statistical data to explain or test a phenomenon. This method typically uses numbers, graphs, tables, or statistical calculations to interpret the data.

Researchers conducted an in-depth analysis of the problem they wanted to solve. In this context, the analysis focused on the need to predict blood bag demand at the Indonesian Red Cross (PMI) Blood Donor Unit (UDD) in Langkat Regency. Researchers explored factors influencing blood demand, such as changes in the number of donors. The research continued by gathering references from various sources, such as journals, online books, and scientific articles discussing the single moving average (SMA) method for predicting blood demand. This literature review helped researchers understand existing theories and identify gaps or deficiencies in previous research that could serve as a basis for this study.

RESULTS AND DISCUSSION

This chapter will present the results of applying the Single Moving Average (SMA) method to predict blood bag demand at the Indonesian Red Cross (PMI) Blood Donor Unit (UDD) in Langkat Regency, based on historical blood bag demand data. The predictions obtained will be compared with actual data to assess the accuracy and effectiveness of the method.

Furthermore, the discussion will focus on interpreting the prediction results, analyzing the method's suitability to the characteristics of the blood bag demand data, and the implications of the research results for blood stock management at the PMI UDD in Langkat Regency. Therefore, this chapter aims to provide a comprehensive overview of the performance of the Single Moving Average (SMA) data mining method in the context of blood bag demand forecasting and the practical benefits it can provide.

Description of Research Data

The data used in this study is the blood bag demand data at the UDD PMI Langkat Regency during the period January 2024 to June 2025. This data was obtained through UDD administrative records which include blood donor activities, thus providing a comprehensive picture of the demand for blood bags in the region.

The data includes the number of requests per month based on blood type (A, B, AB, O). The following is a summary of blood bag demand data during the period, as seen in Table 1.

Table 1. Blood Demand Data Based on Blood Type

No.	Month	Group A	Group B	Group AB	Group O
1	January 2024	95	150	20	187
2	February 2024	85	105	15	150
3	March 2024	55	60	15	90
4	April 2024	105	120	22	220
5	May 2024	110	65	10	145
6	June 2024	110	70	30	250
7	July 2024	52	50	20	125
8	August 2024	85	100	20	190
9	September 2024	145	120	35	130
10	October 2024	170	155	45	205
11	November 24	80	130	30	120
12	Desember 2024	250	230	60	350
13	January 2025	105	125	15	200
14	February 2025	140	235	44	281
15	March 2025	85	105	15	120
16	April 2025	90	155	15	180

17	May 2025	200	90	40	150
18	June 2025	130	50	30	170
Number of Requests		2.092	2.115	481	3.263

Based on the table above, blood demand fluctuates every month, where blood type O consistently has the highest demand with a peak in December 2024 (350 bags), while blood type AB has the lowest demand which is relatively stable between 10 to 60 bags, and this variation is likely influenced by seasonal factors, extraordinary events, or hospital needs.

The highest average monthly demand for blood was found in blood group O, at around 181.28 units, while the lowest was in blood group AB, at around 26.72 units. Meanwhile, blood groups A and B had nearly identical average demands, at around 116.22 and 117.5 units, respectively. This data was then used as the basis for predicting the following month's demand using the Single Moving Average (SMA) method.

Comparison of Prediction Results with Actual Data

After calculating the predicted demand for blood bags using the Single Moving Average (SMA) method, the next step is to compare the predicted results with the actual demand data. Error, Mean Absolute Deviation (MAD), Mean Squared Error (MSE), and Mean Absolute Percentage Error (MAPE) were calculated using RStudio. This comparison was conducted to measure the accuracy of the SMA method in predicting blood bag demand at the Langkat Regency PMI Blood Donor Unit (UDD).

Evaluation of the Accuracy of Blood Type A Demand Prediction Results

To evaluate the accuracy of blood bag demand predictions using the Single Moving Average method, an error value was calculated between the predicted and actual data. This error indicates the magnitude of the difference between the predicted and actual values, and can be used to assess the quality of the forecasting model. The following is a calculation of the error accuracy of the blood group A demand prediction using RStudio, as shown in the following figure.

	Periode	DataAsli	MovingAverage3	Error
1	1	95	NA	NA
2	2	85	NA	NA
3	3	55	NA	NA
4	4	105	78.33333	26.66667
5	5	110	81.66667	28.33333
6	6	110	90.00000	20.00000
7	7	52	108.33333	-56.33333
8	8	85	90.66667	-5.66667
9	9	145	82.33333	62.66667
10	10	170	94.00000	76.00000
11	11	80	133.33333	-53.33333
12	12	250	131.66667	118.33333
13	13	105	166.66667	-61.66667
14	14	140	145.00000	-5.00000
15	15	85	165.00000	-80.00000
16	16	90	110.00000	-20.00000
17	17	200	105.00000	95.00000
18	18	130	125.00000	5.00000
19	19	NA	140.00000	NA
20	Summary	NA	NA	NA
21	MAD	NA	NA	47.60000
22	MSE	NA	NA	3423.17040
23	MAPE (%)	NA	NA	41.08000

Figure 1. Evaluation of the Accuracy of Blood Type A Prediction Results

Evaluation of the Accuracy of Blood Type B Demand Prediction Results

The following is a calculation of the accuracy error of the results of predicting the demand for blood group B bags using RStudio, as seen in the following image.

	Periode	DataAsli	MovingAverage3	Error
1	1	150	NA	NA
2	2	105	NA	NA
3	3	60	NA	NA
4	4	120	105.00000	15.00000
5	5	65	95.00000	-30.00000
6	6	70	81.66667	-11.66667
7	7	50	85.00000	-35.00000
8	8	100	61.66667	38.33333
9	9	120	73.33333	46.66667
10	10	155	90.00000	65.00000
11	11	130	125.00000	5.00000
12	12	230	135.00000	95.00000
13	13	125	171.66667	-46.66667
14	14	235	161.66667	73.33333
15	15	105	196.66667	-91.66667
16	16	155	155.00000	0.00000
17	17	90	165.00000	-75.00000
18	18	50	116.66667	-66.66667
19	19	NA	98.33333	NA
20	Summary	NA	NA	NA
21	MAD	NA	NA	46.33330
22	MSE	NA	NA	3029.07410
23	MAPE (%)	NA	NA	45.48000

Figure 2. Evaluation of the Accuracy of Blood Type B Demand Prediction Results

Evaluation of the Accuracy of the Prediction Results for Blood Type AB Demand

The following is a calculation of the accuracy error of the results of predicting the demand for blood bags for type AB using RStudio, as seen in the following image.

	Periode	DataAsli	MovingAverage3	Error
1	1	20	NA	NA
2	2	15	NA	NA
3	3	15	NA	NA
4	4	22	16.66667	5.333333
5	5	10	17.33333	-7.333333
6	6	30	15.66667	14.333333
7	7	20	20.66667	-0.666667
8	8	20	20.00000	0.000000
9	9	35	23.33333	11.666667
10	10	45	25.00000	20.000000
11	11	30	33.33333	-3.333333
12	12	60	36.66667	23.333333
13	13	15	45.00000	-30.000000
14	14	44	35.00000	9.000000
15	15	15	39.66667	-24.666667
16	16	15	24.66667	-9.666667
17	17	40	24.66667	15.333333
18	18	30	23.33333	6.666667
19	19	NA	28.33333	NA
20	Summary	NA	NA	NA
21	MAD	NA	NA	12.0889000
22	MSE	NA	NA	222.8148000
23	MAPE (%)	NA	NA	52.4200000

Figure 3. Evaluation of the Accuracy of the Prediction Results for Blood Type AB Demand

Evaluation of the Accuracy of Blood Type O Demand Prediction Results

The following is a calculation of the accuracy error of the results of predicting the demand for blood type O bags using RStudio, as seen in the following image.

	Periode	DataAsli	MovingAverage3	Error
1	1	187	NA	NA
2	2	150	NA	NA
3	3	90	NA	NA
4	4	220	142.3333	77.666667
5	5	145	153.3333	-8.333333
6	6	250	151.6667	98.333333
7	7	125	205.0000	-80.000000
8	8	190	173.3333	16.666667
9	9	130	188.3333	-58.333333
10	10	205	148.3333	56.666667
11	11	120	175.0000	-55.000000
12	12	350	151.6667	198.333333
13	13	200	225.0000	-25.000000
14	14	281	223.3333	57.666667
15	15	120	277.0000	-157.000000
16	16	180	200.3333	-20.333333
17	17	150	193.6667	-43.666667
18	18	170	150.0000	20.000000
19	19	NA	166.6667	NA
20	Summary	NA	NA	NA
21	MAD	NA	NA	64.866700
22	MSE	NA	NA	6849.563000
23	MAPE (%)	NA	NA	36.280000

Figure 4. Evaluation of the Accuracy of Blood Type O Demand Prediction Results

CONCLUSION

Based on the results of research on the Application of Data Mining in Predicting Blood Bag Demand Based on Blood Type using the Single Moving Average (SMA) Method Case Study: UDD PMI Langkat Regency, it can be concluded that the SMA method is an effective and practical forecasting technique to be used in this context. This method is able to process historical blood demand data well, producing fairly accurate predictions especially for blood types with relatively stable demand patterns such as blood types A, B, and O. From the analysis results, blood type O showed the highest demand consistently during the observation period, with an average demand reaching 161 bags per month, followed by blood types A and B which also experienced an increasing trend in demand. Blood type AB has the lowest demand but still requires special attention so that stocks do not run out when a sudden spike occurs. The SMA method used for a certain period (3 months) provides prediction results that are close to the actual data with an average Mean Absolute Percentage Error (MAPE) value below 10%, which indicates a level of prediction error that is still within reasonable limits and can be relied upon for blood stock planning. The use of SMA helps the PMI UDD in anticipating blood needs so that it can reduce the risk of stock shortages that could potentially disrupt medical services, as well as avoid excess stock that can cause waste due to expired blood.

REFERENCES

Buku

- [1] Indah Purnama Sari. Algoritma dan Pemrograman. Medan: UMSU Press, 2023, pp. 290.
- [2] Indah Purnama Sari. Buku Ajar Pemrograman Internet Dasar. Medan: UMSU Press, 2022, pp. 300.
- [3] Indah Purnama Sari. Buku Ajar Rekayasa Perangkat Lunak. Medan: UMSU Press, 2021, pp. 228.

- [4] Janner Simarmata Arsan Kumala Jaya, Syarifah Fitrah Ramadhani, Niel Ananto, Abdul Karim, Betrisandi, Muhammad Ilham Alhari, Cucut Susanto, Suardinata, Indah Purnama Sari, Edson Yahuda Putra. *Komputer dan Masyarakat*. Medan: Yayasan Kita Menulis, 2024, pp.162.
- [5] Mahdianta Pandia, Indah Purnama Sari, Alexander Wirapraja Fergie Joanda Kaunang, Syarifah Fitrah Ramadhani Stenly Richard Pungus, Sudirman, Suardinata Jimmy Herawan Moedjahedy, Elly Warni, Debby Erce Sondakh. *Pengantar Bahasa Pemrograman Python*. Medan : Yayasan Kita Menulis, 2024, pp.180
- [6] Zelvi Gustiana Arif Dwinanto, Indah Purnama Sari, Janner Simarmata Mahdianta Pandia, Supriadi Syam, Semmy Wellem Taju Fitrah Eka Susilawati, Asmah Akhriana, Rolly Junius Lontaan Fergie Joanda Kaunang. *Perkembangan Teknologi Informatika*. Medan: Yayasan Kita Menulis, 2024, pp.158

Jurnal

- [7] Sari, I.P., Hariani, P.P., Al-Khowarizmi, A., Ramadhani, F., Sulaiman, O.K., Satria, A., & Manurung, A.A. (2024). CLUSTERING HIV/AIDS DISEASE USING K-MEANS CLUSTERING ALGORITHM. *Proceeding International Seminar on Islamic Studies* 5 (1), 1668-1676
- [8] Sari, I.P., Ramadhani, F., Satria, A., & Sulaiman, O.K. Leukocoria Identification: A 5-Fold Cross Validation CNN and Adaboost Hybrid Approach. *2023 6th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, 486-491
- [9] Manurung, A.A., Nasution, M.D., & Sari, I.P. (2023). Implementation of Fuzzy K-Nearest Neighbor Method in Dengue Disease Classification. *2023 11th International Conference on Cyber and IT Service Management (CITSM)*, 1-4
- [10] Sari, I.P., Ramadhani, F., Satria, A., & Apdilah, D. (2023). Implementasi Pengolahan Citra Digital dalam Pengenalan Wajah menggunakan Algoritma PCA dan Viola Jones. *Hello World Jurnal Ilmu Komputer* 2 (3), 146-157
- [11] Sari, I.P., Al-Khowarizmi, A, Sulaiman, O.K., & Apdilah, D. (2023). Implementation of Data Classification Using K-Means Algorithm in Clustering Stunting Cases. *Journal of Computer Science, Information Technology and Telecommunication Engineering* 4 (2), 402-412
- [12] Sulaiman, O.K & Batubara, I.H. (2021). Implementation Data Mining For Level Analysis Traffic Violation By Algorithm Association Rule. *Al'adzkiya International of Computer Science and Information Technology (AIOCSIT) Journal* 2 (2), 128-135
- [13] Sari, I.P., Batubara, I.H., & Al-Khowarizmi, A. (2021). Sensitivity Of Obtaining Errors In The Combination Of Fuzzy And Neural Networks For Conducting Student Assessment On E-Learning. *International Journal of Economic, Technology and Social Sciences (Injects)* 2 (1), 331-338
- [14] Sari, I.P., Al-Khowarizmi, A., & Batubara, I.H. (2021). Cluster Analysis Using K-Means Algorithm and Fuzzy C-Means Clustering For Grouping Students' Abilities In Online Learning Process. *Journal of Computer Science, Information Technology and Telecommunication Engineering* 2 (1), 139-144
- [15] Apdilah, D., & Sari, I.P. (2021). Optimization Of The Fuzzy C-Means Cluster Center For Credit Data Grouping Using Genetic Algorithms. *Al'adzkiya International of Computer Science and Information Technology (AIOCSIT) Journal* 2 (2), 156-163