

Click here and write your Article Category

Application of Data Mining to Analyze BPJS Patient Satisfaction Levels Regarding the Service Attitude of PTPN II Tanjung Selamat General Hospital Using the K-Means Clustering Algorithm

Yohanni Syahra^{1*}, Natasya Nabaceva²

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

²Department of Information System, STMIK Tri Guna Darma, Medan, North Sumatra, Indonesia

ARTICLE INFORMATION

Received: February 00, 00
Revised: March 00, 00
Available Online: April 00, 00

KEYWORDS

BPJS Health; patient satisfaction; Data Mining; K-Means Clustering

CORRESPONDENCE

Phone: +6281260003950
E-mail: yohannisyahra@umsu.ac.id

A B S T R A C T

The importance of the BPJS Kesehatan's role in providing affordable healthcare access to the community is undeniable. Rumah Sakit Umum PTPN II Tanjung Selamat collaborates with BPJS Kesehatan to deliver services to BPJS participants. However, patient satisfaction levels have been declining, with primary complaints including long registration processes, data recording errors, extended waiting times for doctor consultations, and inadequate hospital facilities. Understanding the factors affecting patient satisfaction is crucial for improving healthcare quality at this hospital. To address these issues, this study applies Data Mining methods, specifically the K-Means Clustering algorithm, to analyze BPJS patient satisfaction levels at Rumah Sakit Umum PTPN II Tanjung Selamat. K-Means Clustering is chosen for its ability to group data based on similar characteristics, allowing the identification of patterns influencing patient satisfaction. Utilizing a desktop application designed with Visual Basic 2010, this research provides flexibility in determining the number of clusters, aiding in more effective analysis and grouping of patient satisfaction data. The findings are expected to offer valuable insights for the management of Rumah Sakit Umum PTPN II Tanjung Selamat to enhance healthcare services for BPJS participants. By identifying patient groups based on their satisfaction levels, the hospital can take more targeted actions to improve unsatisfactory aspects. This study demonstrates that the application of Data Mining with the K-Means method can be an effective tool in evaluating and improving healthcare service quality.

INTRODUCTION

BPJS Kesehatan (Health Social Security Administering Body) plays a crucial role in providing affordable healthcare access to the public [1]. PTPN II Tanjung Selamat General Hospital is one of the healthcare facilities collaborating with BPJS Kesehatan to provide healthcare services to BPJS participants. Patient satisfaction is a key indicator in evaluating the quality of healthcare services, and a thorough understanding of the factors that influence patient satisfaction levels is crucial. This decline is evident from the results of the patient satisfaction questionnaire, which shows that many patients complained about the lengthy registration process and frequent errors in data recording, difficulty in accessing healthcare services such as long waiting times for doctor consultations, and hospital facilities that are considered inadequate, such as uncomfortable waiting rooms and poor hygiene.

Based on the patient satisfaction questionnaire that has been created, several important findings can be identified. For example, in the aspect of patient administration, the questionnaire showed that many patients felt the registration and administrative handling process still took a long time and often made errors in data recording. This caused discomfort

and dissatisfaction among patients. Furthermore, in terms of accessibility, many patients complained about the difficulty in gaining access to necessary health services, such as long waiting times for doctor consultations or difficulty in obtaining information regarding doctor schedules and inpatient room availability. Hospital facilities were also highlighted, with many patients considering facilities such as waiting rooms, cleanliness, and the availability of medical equipment to be still inadequate. Therefore, a solution to these problems can be overcome with a system using the application of Data Mining. Data Mining is the extraction or discovery of new information by searching for certain patterns or rules from a very large amount of data. Data Mining is also referred to as a series of processes to extract added value in the form of knowledge that has not been known manually from a data set. In this study, data mining analysis was carried out using the K-Means method [2]. The K-Means method divides data into several groups so that data with the same characteristics are in the same cluster and data with different characteristics are in different clusters [3].

The purpose of this study is to cluster the level of BPJS patient satisfaction towards the service attitude of PTPN II Tanjung Selamat General Hospital. The K-Means method was chosen because this method must use physical data that is not abstract and is clear, this is in accordance with the data that will be used in the problem in data grouping [4]. In this study, a Desktop Programming-based software will be designed. The desktop application used is Visual Basic 2010 and the K-Means method is flexible because the user can determine the number of clusters to be created.

METHOD

Data Mining

Data mining is part of the Knowledge Discovery in Database (KDD) process [5]. With data mining, we can classify, predict, estimate, and obtain other useful information from large data sets [6]. Data mining is the process of finding interesting patterns or information in selected data using specific techniques or methods. Data mining techniques, methods, or algorithms vary widely. The selection of the appropriate method or algorithm depends heavily on the overall goal and process of Knowledge Discovery in Database (KDD) [7].

Data mining is the process of finding interesting patterns or information in selected data using specific techniques or methods. Data mining techniques, methods, or algorithms vary widely [8]. The selection of the appropriate method or algorithm depends heavily on the overall goal and process of Knowledge Discovery in Database (KDD) [9].

K-Means Method

The K-Means algorithm is a partitional algorithm because it is based on determining the initial number of clusters by defining the initial centroid values. The K-Means algorithm uses an iterative process to obtain a cluster database [10]. It takes the desired initial number of clusters as input and produces the final centroid points as output. The K-Means method randomly selects k clusters as the initial centroid points. The number of iterations to reach the cluster centroids is influenced by the random selection of the initial cluster centroids [11]. Therefore, a method for developing the algorithm is to determine cluster centroids based on high initial data density to achieve higher performance. used the K-Means algorithm in their research to improve the speed of the clustering method by simplifying centroid determination [12].

RESULTS AND DISCUSSION

Application of the K-Means Method

To process the data in Table 1 above using the K-Means algorithm, normalize or clean unused data and initialize the shipping and address data. The following are the steps in the K-Means algorithm, specifying the number of clusters, for example, $k = 3$.

Table 1. Cluster Table

No	Cluster	Information
1	K-1	Very Satisfied
2	K-2	Satisfied
3	K-3	Quite Satisfied

Based on the Knowledge Discovery in Database (KDD) process applied to patient administration at PTPN II Tanjung Selamat General Hospital, it was found that of the total patients analyzed, 30 patients felt very satisfied, with a distribution of 22 males and 8 females; 7 patients were in Class 1, 8 patients in Class 2, and 15 patients in Class 3. A total of 58 patients felt satisfied, with 39 males and 19 females; 16 patients in Class 1, 14 patients in Class 2, and 28 patients in Class 3. Meanwhile, 8 patients felt quite satisfied, consisting of 6 males and 2 females; 1 patient in Class 1, 2 patients in Class 2, and 5 patients in Class 3. These results indicate that most of the patients who felt very satisfied and satisfied were male, and the majority of satisfied patients were in Class 3.

System Implementation

The function of this interface is to provide input and display output from the application. This application has an interface consisting of a login menu, BPJS patient satisfaction data, and a K-Means process menu. The main page displays the menus at the beginning of the system, namely the login menu and the main menu. The main page menus are as follows:

1. Login Menu

The login menu is used to secure the system from unauthorized users before accessing the main menu. The login menu appears as follows:

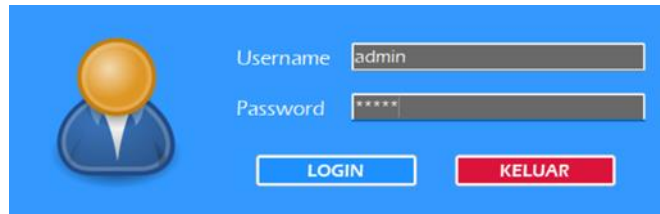


Figure 1. Login Menu

2. Main Menu

The Main Menu serves as a link between the Patient Data, processes, and reports menus. The Main Menu display is shown below.



Figure 2. Main Menu

The administrator displays the data processing menu for storing data in the database, namely the Patient Data Menu. The main admin page menu is as follows:

1. Patient Data Menu

The patient data menu is used for processing data input, data changes, and deletion. The patient data menu is as follows.



Figure 3. Patient Data Menu

2. Centroid Data Menu

The centroid data menu displays centroid data. The patient data menu is as follows.

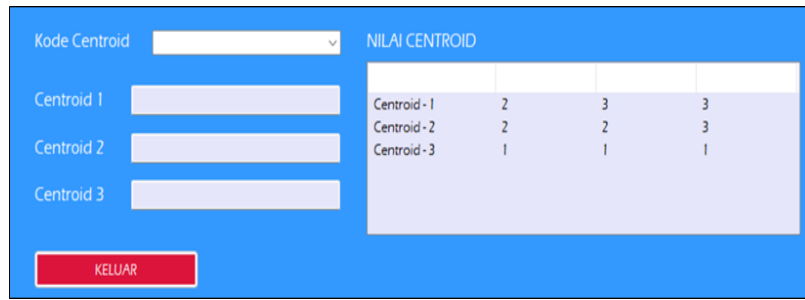


Figure 4. Centroid Data Menu

In this section, we are asked to conduct tests using new data samples to verify the accuracy of the system designed using previously tested and calibrated tools. The results of the calculation process in the patient data grouping program are as follows.



Figure 5. K-Means Analysis Results



Figure 6. Grouping Results Report

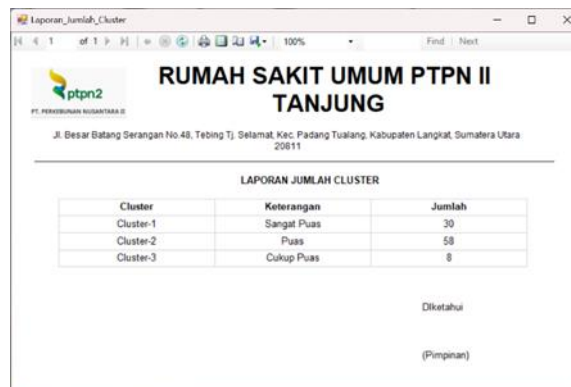


Figure 7. Cluster Number Results Report

CONCLUSION

Based on an analysis of the application of the K-Means Clustering algorithm to group BPJS patient satisfaction levels at PTPN II Tanjung Selamat General Hospital, it can be concluded that this method is effective in identifying and categorizing patient satisfaction levels. By grouping satisfaction data into categories such as Very Satisfied, Satisfied, and Quite Satisfied, the hospital can understand the distribution of patient satisfaction more clearly and structuredly. This allows hospital management to gain better insight into the factors influencing patient satisfaction and take necessary steps to improve service quality. The implementation of the K-Means method in the designed system involves several stages, including determining the number of clusters, randomly selecting an initial centroid, grouping the data based on the closest distance to the centroid, and updating the centroid until convergence is achieved. This process allows for the efficient and accurate identification of key clusters of patient satisfaction levels. The clustering results demonstrate a clear distribution of patient satisfaction levels with hospital services, providing a strong basis for the hospital to make more informed and targeted decisions to improve services. The system built to analyze patient satisfaction levels using the K-Means Clustering algorithm has proven effective and reliable. Built using the Visual Basic programming language and Microsoft Access database, and designed with UML, this system provides a user-friendly interface that makes it easy for admins to manage and analyze patient satisfaction data. System testing shows that features such as the login form, main menu, patient satisfaction data processing, and clustering process function well, proving that this system can be relied upon to analyze patient satisfaction levels and provide valuable information for hospitals to improve the quality of healthcare services.

REFERENCES

Book: Single Author

- [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.

Book: Two or More Authors

- [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
- [7] Wahyuddin S, Indah Purnama Sari, Samuel Hasudungan Tampubolon Andi Zulherry, Butsiarah, Taufik, George Morris William Tangka Rahmad Bahri, Mhd. Basri, Raemon Syaljumairi, Robbi Rahim Muharman Lubis, Baso Ali, Sirmayanti, Janner Simarmata. *Cloud Computing: Referensi Lengkap*. Medan: Yayasan Kita Menulis, 2025, pp.248.
- [8] Muharman Lubis Ilham Firman Ashari, Debby Erce Sondakh, Rahmawati Rolly Junius Lontaan, Mustarum Musaruddin Indah Purnama Sari, Muh. Nadzirin Anshari Nur, Hanalde Andre Muh. Rais, Janner Simarmata. *Internet of Things (IoT) Dan Multimedia: Integrasi Dan Aplikasi*. Medan: Yayasan Kita Menulis, 2024, pp.182

Journal Article from a Full Text Database

- [9] E. Faturahmah and B. B. Raharjo, "Analisis Tingkat Kepuasan Pasien BPJS Terhadap Kualitas Pelayanan Petugas Medis Medis Melalui Faktor Assurance di Ruang Rawat Inap RSUD Bima, NTB," *Public Health Perspective Journal*, vol. II, no. 3, pp. 254 - 261, 2017.
- [10] F. Yunita, "PENERAPAN DATA MINING MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING PADA PENERIMAAN MAHASISWA BARU (STUDI KASUS : UNIVERSITAS ISLAM INDRAGIRI)," 2018.
- [11] Zulherry, A., Riadi, I., & Umar, R. (2026). Anomaly Detection in Cloud Device-Based Information Technology Infrastructure Using Isolation Forest Algorithm. *Journal Of Informatics And Telecommunication Engineering* 9 (2)
- [12] Sari, I.P., & Zulherry, A. (2025). Development of A Smart Monitoring System for IoT-Based Tide Observation. *Al'adzkiya International of Computer Science and Information Technology (AIOCSIT) Journal*, 6 (2), 18-23

- [13] Zulherry, A., Gunawan, M., & Sari, I.P. (2025). Development of an Android-Based Smart Health Monitoring Device for Heartbeat Detection. *Al'adzkiya International of Computer Science and Information Technology (AloCSIT) Journal*, 6 (2), 43-47
- [14] Sari, I.P., Zulherry, A., Basri, M., & Hayani W. (2025). Pembelajaran Pemrograman berbasis Machine Learning sebagai Upaya Peningkatan Computational Thinking. *Jurnal Penelitian, Pendidikan dan Pengajaran: JPPP* 6 (3), 245-250
- [15] T. Amalina, D. B. A. Pramana and B. N. Sari, "Metode K-Means Clustering Dalam Pengelompokan Penjualan Produk Frozen Food," *Jurnal Ilmiah Wahana Pendidikan*, vol. VIII, pp. 574-583, 2022.
- [16] S. Butsianto and N. T. Mayangwulan, "Penerapan Data Mining Untuk Prediksi Penjualan Mobil Menggunakan Metode K-Means Clustering," *Jurnal Nasional Komputasi dan Teknologi Informasi*, vol. III, no. 3, 2020.
- [17] Bisono, A. T., & Zulherry, A. (2025). Analisis sentimen game Genshin Impact untuk mengetahui reaksi dan harapan pemain menggunakan metode Naïve Bayes. *sudo Jurnal Teknik Informatika*, 4(2), 183-193.
- [18] Basri, M., & Zulherry, A. (2025). Analysis of the Impact of Gambling and Online Loans in the Perspective of Informatics, Islam, and Kemuhammadiyah. *Ar-Rasyid: Jurnal Pendidikan Agama Islam*, 5(1), 65-73.
- [19] Asadel, A., & Zulherry, A. (2025). Detecting Zero-Width Characters Obfuscated in Phishing URLs using the XGBOOST Algorithm. *Hanif Journal of Information Systems* 3 (1), 43-53
- [20] N. S. Nurajizah and A. Salbinda, "Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Pada Toko Fashion Hijab Banten," *Jurnal Teknik Komputer AMIK BSI*, vol. VII, no. 2, pp. 158-163, 2021.
- [21] A. Nugraha, O. Nurdiawan and G. Dwilestari, "Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Pada Toko Yana Sport," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. VI, no. 2, pp. 849-855, 2022.
- [22] 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
- [23] Zulherry, A., Sari, I.P., & Basri, M. (2025). Perancangan Aplikasi Monitoring Kehadiran Pegawai Menggunakan RFID. *sudo Jurnal Teknik Informatika* 4 (4), 378-384
- [24] Ichsan, A., Zulherry, A., Lubis, T. A., & Shahnaz, B. A. Z. (2025). Utilization of Mobile Applications to Speed Up The Search for Android-Based Index Places. *IJATCoS: Indonesian Journal of Applied Technology. Computer and Science*, 2(1).
- [25] Amada, E.P., & Zulherry, A. (2025). Klasterisasi Minat Dan Bakat Siswa Menggunakan Metode X-Means Berbasis Web: Studi Kasus SMA Negeri 1 Hamparan Perak. *sudo Jurnal Teknik Informatika* 4 (4), 276-283.
- [26] Zulherry, A., Ramadhani, F., & Satria, A. (2024). Klasifikasi Data Tracer Study Dengan Pemanfaatan Data Mining Menggunakan Algoritma Support Vector Machine dan Neural Network. *Portal Riset dan Inovasi Sistem Perangkat Lunak* 2 (1), 45-54
- [27] Zulherry, A. (2023). Decision making for network security with simple additive weighting method. *Journal of Intelligent Decision Support System (IDSS)*, 6(3), 155-159.
- [28] Zulherry, A., Gunawan, T. S., & Wanayumini, W. (2021). Analisis Hasil Pendukung Keputusan Mendapatkan Rumah Dinas Perusahaan Menggunakan Metode Analytical Hierarchy Process (AHP) dan Technique for Order Preference by Similarity to Ideal Solution (TOPSIS). *Jurnal Media Informatika Budidarma*, 5(2), 695-704.
- [29] 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
- [30] 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
- [31] Y. Mardi, "Data Mining : Klasifikasi Menggunakan Algoritma C4.5," *Jurnal Edik Informatika*, vol. V, no. 2407-0491, pp. 213-219, 2018.
- [32] J. Nasir, "Penerapan Data Mining Clustering Dalam Mengelompokan Buku Dengan Metode K-Means," *Jurnal SIMETRIS*, vol. XI, no. 2, pp. 1-13, 2020.
- [33] T. Amalina, D. B. A. Pramana and B. N. Sari, "Metode K-Means Clustering Dalam Pengelompokan Penjualan Produk Frozen Food," *Jurnal Ilmiah Wahana Pendidikan*, vol. VIII, no. 15, pp. 574-583, 2022.
- [34] S. Handoko, F. and E. T. E. Handayani, "Implementasi Data Mining Untuk Menentukan Tingkat Penjualan Paket Data Telkomsel Menggunakan Metode K-Means Clustering," *Jurnal Ilmiah Teknologi dan Rekayasa*, vol. XXV, no. 1, pp. 76-88, 2020.
- [35] 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

- [36] 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
- [37] 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
- [38] 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
- [39] 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
- [40] 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
- [41] M. R. Alhapizi, M. Nasir and I. Effendy, "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Mahasiswa Baru Universitas Bina Darma Palembang," *Journal of Software Engineering Ampera*, vol. I, no. 1, pp. 1-14, 2020.
- [42] F. Dikarya and S. Muharni, "Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Universitas Terbaik Di Dunia," *Jurnal Informatika*, vol. XXII, no. 2, pp. 124-131, 2022.
- [43] R. A. Indraputra and R. Fitriana, "K-Means Clustering Data COVID-19," *Jurnal Teknik Industri*, vol. X, no. 2622-5131, pp. 275-282, 2020.