

CLASSIFICATION OF PATIENT SATISFACTION LEVEL ON HEALTH SERVICES USING THE C4.5 ALGORITHM

Sriani¹, Aidil Halim Lubis², Sofiah³

Computer Science
North Sumatera State Islamic University
Medan, Indonesia
sriani@uinsu.ac.id, aidilhalimlubis@uinsu.ac.id, sofiah0701193155@uinsu.ac.id

(*) Corresponding Author

Abstract

Quality health services are related to patient satisfaction. Patient satisfaction can be used as a benchmark for improving the quality of health services. Problems often occur when implementing health techniques, such as service problems at the Restu Clinic. Patients and their families indicated that although the Restu Clinic was established with adequate facilities, it had not yet reached the maximum level of service. These indications include long waiting times for examinations, a lack of thoroughness by medical personnel, and services that are not timely. Service quality cannot be separated from the dimensions that are the core of quality services, which are expected to meet patient needs. Patient satisfaction is considered an important indicator of good quality. This research will only discuss four aspects of service quality, which are reliability, responsiveness, assurance, and empathy, from health workers at the Restu Clinic. The C4.5 algorithm is known to be superior in producing decision trees that efficiently solve discrete and numerical variables and provide satisfactory accuracy. Therefore, the author conducted a study to assess service quality using the C4.5 algorithm. This research aims to determine the factors that influence the quality of health services and to know patient satisfaction with health services at the Restu Clinic. Knowing the intensity of patient satisfaction with services at the Restu Clinic can improve the quality of optimal services and gain patients' trust in government agencies.

Keywords: Patient satisfaction level; C4.5 algorithm; health services.

Abstrak

Pelayanan kesehatan yang berkualitas berkaitan dengan kepuasan pasien. Kepuasan pasien dapat dijadikan tolak ukur untuk meningkatkan kualitas pelayanan kesehatan. Permasalahan sering terjadi pada saat pelaksanaan teknik kesehatan, seperti permasalahan pelayanan di Klinik Restu. Pasien dan keluarga pasien mengindikasikan bahwa meskipun Klinik Restu didirikan dengan fasilitas yang memadai, namun belum mencapai tingkat pelayanan yang maksimal. Indikasi tersebut antara lain waktu tunggu pemeriksaan yang lama, kurangnya ketelitian tenaga medis, dan pelayanan yang tidak tepat waktu. Kualitas pelayanan tidak terlepas dari dimensi-dimensi yang menjadi inti dari pelayanan yang berkualitas, yang diharapkan dapat memenuhi kebutuhan pasien. Kepuasan pasien dianggap sebagai indikator penting dari kualitas yang baik. Penelitian ini hanya akan membahas empat aspek kualitas pelayanan, yaitu kehandalan, daya tanggap, jaminan, dan empati, dari tenaga kesehatan di Klinik Restu. Algoritma C4.5 dikenal unggul dalam menghasilkan pohon keputusan yang secara efisien menyelesaikan variabel diskrit dan numerik serta memberikan akurasi yang memuaskan. Oleh karena itu, penulis melakukan penelitian untuk menilai kualitas pelayanan dengan menggunakan algoritma C4.5. Penelitian ini bertujuan untuk mengetahui faktor-faktor yang mempengaruhi kualitas pelayanan kesehatan dan mengetahui tingkat kepuasan pasien terhadap pelayanan kesehatan di Klinik Restu. Dengan mengetahui intensitas kepuasan pasien terhadap pelayanan di Klinik Restu dapat meningkatkan kualitas pelayanan yang optimal dan mendapatkan kepercayaan pasien terhadap instansi pemerintah.

Kata Kunci: Tingkat kepuasan pasien; Algoritma C4.5; pelayanan kesehatan.

INTRODUCTION

Health is a basic need for every individual. Without good health conditions, a person cannot fulfill his needs (Manzoor, Wei, Hussain, Asif, & Shah, 2019). Therefore, health is every person's capital to live a decent life. Currently, health problems have become an essential need for society. As the standard of living increases, people expect an increase in the quality of their health. Quality health services are health services that pay attention to and prioritize customer needs, hopes, and values. This is an important reference that must be met to provide satisfaction to the community as users of health services (Irmawati, Sultan M, & Nurhannis, 2019)

Health facilities are facilities that Indonesian people need. Clinics are one health service facility implemented to ensure easy, affordable, and quality access to health services to improve public health. The clinic provides basic and special medical services to meet individual needs regarding promotion, prevention, treatment, and recovery based on the Republic of Indonesia Minister of Health Regulation Number 9 of 2014 (Ratnasari & Puspani, 2019).

Problems often occur when implementing health techniques, such as service problems at the Restu Clinic. Patients and their families indicated that although the Restu Clinic was established with adequate facilities, it had not yet reached the maximum level of service (Manzoor et al., 2019). These indications include long waiting times for examinations (Maulana, Tamrin, Alim, & Imran, 2019), a lack of thoroughness by medical personnel (Haqim & Monica, 2021), and services that are not timely (Nur'aeni, Simanjorang, & ., 2020; Sistem et al., 2022). Service quality cannot be separated from the dimensions that are the core of quality services, which are expected to meet patient needs. Patient satisfaction is considered an important indicator of good quality (Ramadhan et al., 2020). The intensity of patient satisfaction has a strong relationship with the rate of patient return, so that it can act as an indicator of service quality.

Based on previous research that has applied the C4.5 algorithm (Aditya Nugroho & Kristiana, 2022). with the research title "Application of the C4.5 Algorithm for Customer Satisfaction of the Chantik Perfume Online Store," the data used is customer satisfaction data with the accuracy results of the C4 algorithm model. 5 shows an accuracy rate of 71%, meaning that the graduation classification model using the C4.5 algorithm is proven to be quite good. The next researcher (Damanik, Wanto, & Gunawan, 2022). has the title "Application of the

C4.5 Decision Tree Algorithm for Classification of Family Welfare Levels in Tiga Dolok Village." The data is based on a survey conducted in Nagori Tiga Dolok, and questionnaires were distributed to the village community of Three Doloks. This research shows that this problem can be overcome by having a decision tree based on predetermined criteria. This research will only discuss four aspects of service quality, which are reliability, responsiveness, assurance, and empathy, from health workers at the Restu Clinic.

Reliability is the ability of health facilities to provide reliable services and meet patient expectations. Responsiveness is the responsiveness of health officers in providing assistance and fast services by conveying information that is easy to understand. Assurance is health workers' attitude, knowledge, and ability to convince patients of health services. Empathy is the attitude of health workers toward paying attention to understanding individual patient needs. Tangible includes the ability of health service facilities to demonstrate their existence to external parties (Effendi, 2020).

Data mining is used to analyze data with statistical applications to obtain information. Data mining can convert large amounts of data into other useful information (Nasrullah 2021). Data mining, the information extraction step from large amounts of data, can be done using various methods and algorithms. Data mining includes methods of classification, association, clustering, prediction, and estimation (Alkhairi & Situmorang, 2022).

Data mining refers to searching for previously unknown information from large amounts of data. In searching for important information by analyzing certain patterns or relationships in large amounts of data(Azwanti 2018), processing data in data mining requires algorithms for extracting information. Algorithm applications in data mining are classified based on their respective functions (Suntoro, 2019).

Decision trees, better known as decision trees, are a classification technique in data mining that is useful for creating tree structures to produce answers to problems that have been entered. A decision tree will make it easier to identify the relationship between the causes of problems and find solutions to problems by considering the causes. With the resulting decision tree structure, it is possible also to analyze it so that new information can be used to solve problems and indirectly provide the best ideas or policies for an organization (Rufiyanto, Rochcham, & Rohman, 2021).

Decision trees have the main benefits, including simplifying decisions that were initially

complex into simple ones and making decisions that tend to interpret problem-solving. Decision trees are useful for investigating data and identifying relationships between input and target variables (Jollyta, Ramdhan, & Zarlis, 2020).

In data mining, The C4.5 algorithm plays a role in classifying decision trees, which are used to make decisions (predictors, univariate/multivariate) (Marisa, Maukar, & Akhriza, 2021). This algorithm is a development of the ID3 algorithm (Decision Tree Induction). The limitation of the ID3 algorithm is that it only works with categorical-type data and cannot be used with numeric-type data. The development of the C4.5 algorithm lies in handling numeric type features, pruning (pruning) the decision tree, and lowering (deriving) the rule set. In addition, the C4.5 algorithm establishes a gain criterion to determine the dividing feature nodes on the resulting tree (Muhamad, Windarto, & Suhada, 2019). The C4.5 algorithm has the main benefit of stopping the process of making complex decisions to become simpler so that decisions tend to interpret solutions to problems (Nasrullah, 2021).

This research was conducted using questionnaire data given to Restu Clinic patients to determine the intensity of patient satisfaction and the factors influencing the quality of service to produce high patient satisfaction with a minimum range of 85%. Knowing the factors that influence the intensity of patient satisfaction can be used as a reference for maximizing optimal service activities for patients to obtain increased service quality.

RESEARCH METHODS

Types of research

This research is quantitative in the context of the data mining field by applying the data mining classification algorithm, namely the C4.5 algorithm (Chicho, Abdulazeez, Zeebaree, & Zebari, 2021; Xue, Qian, Lijuan, Yunchun, & Zheng, 2023). This research focuses on classifying patient satisfaction with health services and determining the accuracy of decision trees. This study aims to apply the C4.5 algorithm to classify the level of patient satisfaction with health services at the Restu Clinic and determine the accuracy of the decision tree in classifying patient satisfaction with health services at the Restu Clinic using the C4.5 algorithm. Implementing the C4.5 algorithm can be done using a data mining application to build a decision tree that can be used to determine the algorithm model's performance so that the decision tree's accuracy can be determined.

Time and Place of Research

Data is collected to determine patient satisfaction with health services by implementing the C4.5 algorithm at Restu Clinic, located on Jl. Simpang Ajamu, Tanjung Sarang Village Elang, Panai Hulu District, Labuhanbatu Regency, North Sumatra 21471, Indonesia. This research was carried out from February 2023 until its completion. This research was 5 stages carried out: planning, data collection, data analysis, model testing, and results.

This study aims to apply the C4.5 algorithm to classify the level of patient satisfaction with health services at the Restu Clinic and determine the accuracy of the decision tree in classifying patient satisfaction with health services at the Restu Clinic using the C4.5 algorithm.

Research Target / Subject

The results of this research show that we can create a classification model for the level of patient satisfaction with health services using the C4.5 algorithm as a decision tree and determine the factors that influence the quality of service at the Restu Clinic.

Procedure

The research framework is the stages used to direct the research conducted, the following Figure 1 explains the stages in this study.

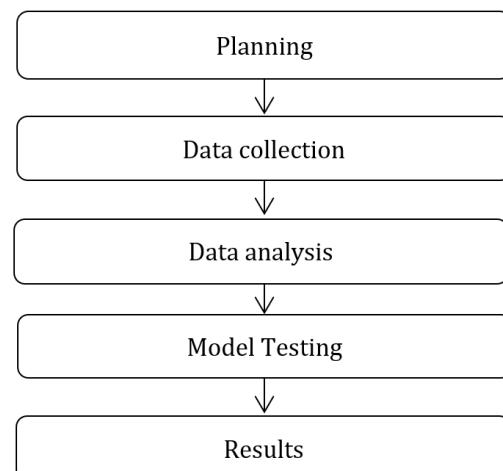


Figure 1. Research Framework

1. Planning

Before conducting research, researchers carry out planning stages, namely identifying and formulating problems, processing research permits, creating a list of questions (questionnaires), collecting supporting tools, analyzing data, and preparing research.

RESULTS AND DISCUSSION

Data analysis

The data needed in this research are the results of the questionnaires given to patients at the Restu clinic. With the target attribute/label "uas" and "It's okay," The basic questions of the questionnaire are based on several references obtained, such as journals published by Adelia et al. (2019).

Algorithm Classification Model C4.5

This classification model will be calculated using respondent data points. To make tree decisions from the data using the C4.5 algorithm classification method, the algorithm calculation uses Microsoft Excel software based on the formula contained in the C4.5 algorithm, which has been explained previously. Figure 3 is the final result of the decision tree obtained by applying a data mining application.

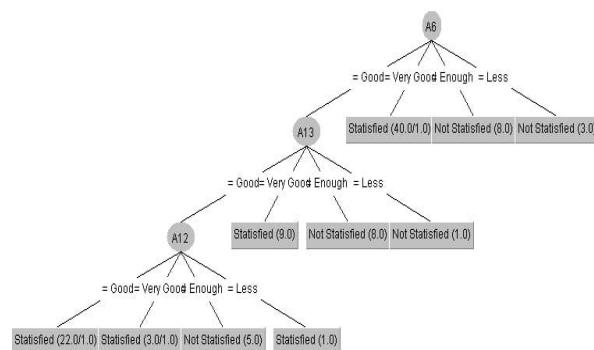


Figure 3. Decision Tree in data mining application

Based on the decision tree above, obtain several rules for classifying patient satisfaction with health services in Restu clinics, including the following:

1. If attribute A6 is rated "very good," then the decision is "good."
2. If attribute A6 is "fair," the decision is "unsatisfactory."
3. If attribute A6 is "poor," the decision is "unsatisfactory."
4. If attribute A6 is "good" and A13 is "very good," then the decision is "good."
5. If attribute A13 is "fair," then the decision is "unsatisfactory."
6. If attribute A13 is rated "poor," then the decision is "unsatisfactory."
7. If attribute A13 is "good" and A12 is "very good," then the decision is "good."

8. If attribute A12 is "good," then the decision is "good."
9. If attribute A12 is "fair," then the decision is "unsatisfactory."
10. If attribute A12 is rated "poor," then the decision is "unsatisfactory."

Data Mining Process

In the previous discussion, the stages of data mining were explained using the C4.5 classification algorithm to create a decision tree using Microsoft Excel and a data mining application. Next, the implementation process will be explained using data mining with the programming language. Then, you will see a decision tree, as shown in Figure 3.

This research has input attributes A1-A20 and target/label attributes, namely "uas" and "not uas." Implementation of the C4.5, found in Chrome. Data mining involves a series of processes for extracting data that has not been discovered. The initial data mining stage starts with raw data to produce information or data already available and processed.

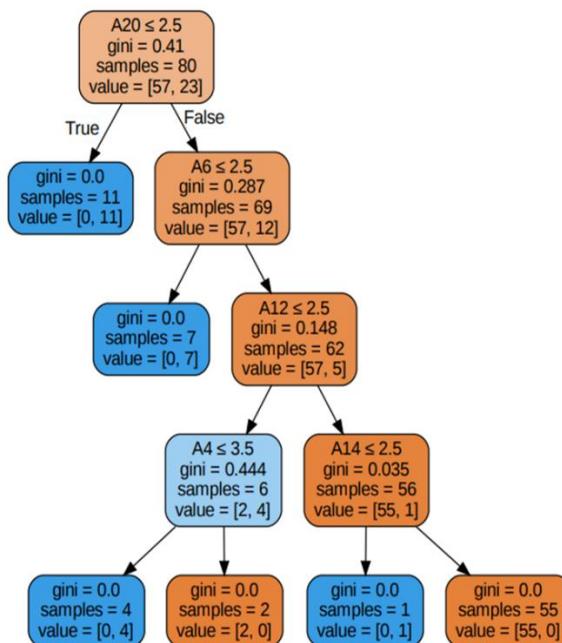


Figure 3. Decision Tree Prediction

Data Cleaning, dirty, or incomplete data causes bad results and system failures that cost time and money. Inappropriate data, such as errors, incomplete, and so on, will be discarded.

Selection is selecting and retrieving data appropriate to the analysis and obtained from the existing data collection.

Transformation is the process of changing the data used into other forms so that the data supports mining procedures.

Data mining is a significant stage involving various techniques for extracting potential patterns to get useful data.

Pattern evolution is the process of finding interesting information by identifying the standards that have been provided. Knowledge presentation is a stage that uses visualization techniques to help users explain the visualization of the analysis results.

Results of the Classification of Patient Satisfaction Levels

To calculate the classification of patient satisfaction, the Restu clinic uses programming language and carries out the coding stage by continuing the program explained first.

The Confusion Matrix is useful for evaluating the work results of supervised learning classification algorithms by paying attention to the accuracy, precision, recall, and F1 score values. The C4.5 algorithm is an algorithm that can be used to make decision trees (Kurniawan, 2019). In measuring the results of the work, four terms describe classification results: TP (*true positive*), FP (*false positive*), TN (*true negative*), and FN (*false negative*) (Ikram, 2023).

Table 1. Patient Satisfaction Level

No	Response	Percentage Level
1	Satisfied	71%
2	Unsatisfied	29%

Based on program codes to classify patient satisfaction at Restu Clinic, patient satisfaction results show that 71% of patient data said they were satisfied with Restu Clinic's services. And 29% of patients said they were unsatisfied with Restu Clinic's services. So the level can be known.

In implementing the C4.5 algorithm using the programming language to classify patient satisfaction levels, the services at the Restu clinic produce very good performance, namely a classification accuracy of 97%.

CONCLUSIONS AND SUGGESTIONS

Conclusion

Based on the analysis that has been carried out using the C4.5 algorithm, information can be used as a reference for improving the quality of health services at the Restu clinic. From the use of the classification method with the C4.5 algorithm, which uses data from 100 respondents, it can be concluded that the level of patient satisfaction is as high as 71%. And 29% were dissatisfied with the services at the Restu Clinic. This classification shows very good performance, namely, an accuracy of 97%.

Application decision tree The C4.5 algorithm with the Python programming language has been able to classify the level of patient satisfaction with Restu Clinic services by obtaining an accuracy rate of 85%. Based on the applications of the C4.5 algorithm, it shows that patient satisfaction greatly influences doctors and health workers responding quickly to patient complaints (A6), as seen from the decision tree. The practical implications of this research provide recommendations for Restu Clinic to improve their services based on the research findings. This shows that this research not only provides academic insights, but also has a direct impact on relevant clinic practices.

Suggestion

The author's suggestion for developing research is To carry out the development of this research, and you can use other data mining methods to make comparisons with other data mining methods and further research regarding the attributes that will later be selected. The findings from this research can be a consideration for the Restu Clinic to improve the quality of their services.

REFERENCES

Aditya Nugroho, C. R., & Kristiana, T. (2022). Penerapan Algoritma C4.5 Untuk Kepuasan Pelanggan Toko Online Parfume Chantik. *Jurnal Algoritme*, 3(1), 10-21. <https://doi.org/10.35957/algoritme.v3i1.3169>

Alkhairi, P., & Situmorang, Z. (2022). Penerapan Data Mining Untuk Menganalisis Kepuasan Pegawai Terhadap Pelayanan Bidang SDM dengan Algoritma C4.5. *Jurasik (Jurnal Riset Sistem Informasi Dan Teknik Informatika)*, 7(1), 40. <https://doi.org/10.30645/jurasik.v7i1.414>



Chicho, B. T., Abdulazeez, A. M., Zeebaree, D. Q., & Zebari, D. A. (2021). Machine Learning Classifiers Based Classification For IRIS Recognition. *Qubahan Academic Journal*, 1(2), 106-118. <https://doi.org/10.48161/QAJ.V1N2A48>

Damanik, S. F., Wanto, A., & Gunawan, I. (2022). Penerapan Algoritma Decision Tree C4.5 untuk Klasifikasi Tingkat Kesejahteraan Keluarga pada Desa Tiga Dolok. *Jurnal Krisnadana*, 1(2), 21-32. <https://doi.org/10.58982/krisnadana.v1i2.108>

Effendi, K. (2020). Tingkat Kepuasan Pasien Terhadap Pelayanan Kesehatan Di Uptd Puskesmas Mutiara Tahun 2019. *Excellent Midwifery Journal*, 3(2), 82-90. <https://doi.org/10.55541/emj.v3i2.127>

Ginting, V. S., Kusrini, K., & Luthfi, E. T. (2020). Penerapan Algoritma C4.5 Dalam Memprediksi Keterlambatan Pembayaran Uang Sekolah Menggunakan Python. *Jurnal Teknologi Informasi*, 4(1), 1-6. <https://doi.org/10.36294/jurti.v4i1.1101>

Haqim, L. N., & Monica, R. D. (2021). Analisis Kuantitatif Pengisian Rekam Medis Pasien Poli Umum Guna Menunjang Mutu Rekam Medis Di Puskesmas Cimahi Tengah. *Jurnal TEDC*, 15(2), 211-215. Retrieved from <http://ejournal.poltekdedc.ac.id/index.php/tecd/article/view/493>

Irmawati, S., Sultan M, H., & Nurhannis, N. (2019). Kualitas Pelayanan Kesehatan di Puskesmas Kecamatan Tatanga Kota Palu. *Jurnal Katalogis*, 5(1), 188-197. Retrieved from <http://jurnal.untad.ac.id/jurnal/index.php/Katalogis/article/view/7968>

Jollyta, D., Ramdhan, W., & Zarlis, M. (2020). *Konsep Data Mining Dan Penerapan*. Yogjakarta: Deepublish.

Kurniawan, A. (2019). Implementasi data mining algoritma c4. 5 untuk memprediksi kelulusan uji kompetensi smk teknik komputer dan jaringan (tkj)(study kasus: smk pembangunan daerah *Informasi Dan Teknologi Ilmiah (INTI)*, 7(1), 5-13. Retrieved from <http://ejurnal.stmik-budidarma.ac.id/index.php/inti/article/view/1801>

Manzoor, F., Wei, L., Hussain, A., Asif, M., & Shah, S. I. A. (2019). Patient Satisfaction with Health Care Services; An Application of Physician's Behavior as a Moderator. *International Journal of Environmental Research and Public Health*, 16(18).

<https://doi.org/10.3390/IJERPH16183318>

Marisa, F., Maukar, A. L., & Akhriza, T. M. (2021). *Data Mining Konsep Dan Penerapannya*. Yogjakarta: Deepublish.

Maulana, D., Tamrin, R., Alim, A., & Imran, A. (2019). Analisis Hubungan Waktu Tunggu Terhadap Kepuasan Pasien Pada Puskesmas Macinni Sombala. *Jurnal Kesehatan*, 12(2), 99-111. <https://doi.org/10.24252/kesehatan.v12i2.10483>

Muhammad, M., Windarto, A. P., & Suhada, S. (2019). Penerapan Algoritma C4.5 Pada Klasifikasi Potensi Siswa Drop Out. *KOMIK (Konferensi Nasional Teknologi Informasi Dan Komputer)*, 3(1), 1-8. <https://doi.org/10.30865/komik.v3i1.1688>

Nasrullah, A. H. (2021). Implementasi Algoritma Decision Tree Untuk Klasifikasi Produk Laris. *Jurnal Ilmiah Ilmu Komputer*, 7(2), 45-51. <https://doi.org/10.35329/jiik.v7i2.203>

Nur'aeni, R., Simanjorang, A., & J. (2020). Pengaruh Mutu Pelayanan Terhadap Kepuasan Pasien Rawat Inap Di Rumah Sakit Izza Karawang. *Journal of Healthcare Technology and Medicine*, 6(2), 1097-1112. <https://doi.org/10.33143/jhtm.v6i2.1152>

Ramadhan, G., Perdana Windarto, A., Irawan, E., Saputra, W., Okprana, H., Infomasi, S., ... Tunas, A. (2020). Penerapan Data Mining Menggunakan Algoritma C4.5 Dalam Mengukur Tingkat Kepuasan Pasien BPJS. *Prosiding Seminar Nasional Riset Dan Information Science (SENARIS)*, 2, 376-385. Pematang Siantar: STIKOM Tunas Bangsa. Retrieved from <http://tunasbangsa.ac.id/seminar/index.php/senaris/article/view/185>

Ratnasari, I., & Puspani, N. S. (2019). Analisis Tingkat Kepuasan Pasien Di Klinik Pratama Abc Kota Bandung Menggunakan Metode Importance Performance Analysis. *Seminar Dan Konferensi Nasional IDEC*, 1-10. Surakarta: Universitas Sebelas Maret. Retrieved from <https://idec.ft.uns.ac.id/wp-content/uploads/2019/05/ID011.pdf>

Rufiyanto, A., Rochcham, M., & Rohman, A. (2021). *Penerapan Algoritma C4.5 Untuk Prediksi Kepuasan Mahasiswa Tahun 2020*. Yogjakarta: Deepublish.

Sistem, A., Pendaftaran, P., Poliklinik, O., Di, J., Sakit, R., Pusat, U., ... Fazary, A. (2022). Analisis Sistem Pelayanan Pendaftaran Online Poliklinik Jantung di Rumah Sakit Umum Pusat Dr. Sardjito Yogyakarta. *Multidisciplinary Journal*, 5(2), 32-40.



<https://doi.org/10.19184/MULTIJOURNAL.V5I2.43557>

Suntoro, J. (2019). *Data Mining: Algoritma dan Implementasi dengan Pemrograman PHP*. Jakarta: Elex Media Komputindo.

Xue, Y., Qian, Y., Lijuan, Z., Yunchun, M., & Zheng, W. (2023). Nursing Quality Based on Decision Tree Algorithm. *2023 IEEE International Conference on Electrical, Automation and Computer Engineering (ICEACE)*, 657–661. <https://doi.org/10.1109/ICEACE60673.2023.10442352>

