JINAV: Journal of Information and Visualization *ISSN: 2746-1440 (online)*
Vol. 3 No. 2 (2022)
https://doi.org/10.35877/454RI.jinav1410

Classification in the Self Monitoring System for Chronic Kidney Failure Patients on Hemodialysis Therapy with SVM

Teguh Andriyanto* & Erna Daniati

Universitas Nusantara PGRI Kediri, Kediri 64112, Indonesia

Abstract

Patients with chronic kidney failure (CKF) need intensive care or therapy. Chronic kidney failure is a condition when kidney function decreases gradually due to damage to kidney tissue. Patients with these conditions need to undergo hemodialysis therapy which can be done every week. The results of this therapy need to be monitored to determine the quality and action after therapy. Monitoring of therapeutic results was initially carried out with medical records where the files were kept by the health agency. In addition, patients need to consult a doctor or nurse to read this medical record. This becomes an obstacle for patients to know the progress of their therapy results. Therefore it is necessary to have a monitoring system as a management information system for patients with chronic kidney failure. Conventional monitoring systems need to include visualization methods or quality fees from doctors. Therefore it is necessary to include the classification method as knowledge of the system to study patterns or rules from doctors in handling the quality of therapeutic results. This classification method uses the Support Vector Machine (SVM) with the Linear kernel as the classification method. In this study, the accuracy obtained with SVM was 90% with a short data training processing time.

Keywords: monitoring system, classification, svm, chronic kidney failure.

1. Introduction

Chronic kidney failure (CKF) is a serious disease that has an increasing number of cases each year (Anonim, 2017). This is what causes CKF to become a deadly disease. Initially, CKF patients do not realize that they have CKF. Patients will experience initial symptoms, such as weakness, nausea, decreased appetite and weight loss which will begin to be felt for several months. According to Basic Health Research (Riskesdas) data for 2013, the prevalence of the Indonesian population suffering from Kidney Failure is 0.2% or 2 per 1000 population and the prevalence of Kidney Stones is 0.6% or 6 per 1000 population (Anonim, 2017). The highest prevalence of CKF is in Central Sulawesi Province, which is 0.5%. The increase in the burden of health costs for services of catastrophic diseases has also increased based on data from the 2016 Ministry of Health Financing and Health Insurance Center. .3 trillion. CKF is a catastrophic disease with high costs and complications can endanger sufferers who experience it.

Until now, there is still no medicine for CKF patients. So that CKF patients must undergo kidney transplantation, Continuous Ambulatory Peritoneal Dialysis (CAPD) therapy or Hemodialysis (HD) therapy to maintain life (Triaji et al., 2017). Initially, patients with CKF on HD therapy will usually feel confused about the condition and the new habits that must be followed. CKF patients on HD therapy should undergo HD twice a week for 4-5 hours. For 2 new patients with CKF HD, they will usually feel confused to know their body condition which is still in the adjustment stage. This is related to the nutritional and fluid diet that must be done. Besides that, HD CKF patients also lack information about CKF. This is one of the reasons why many CKF patients experience death. Based on data from the Pusdatin Ministry of Health in 2015, there were 1,243 deaths in patients undergoing hemodialysis during 2015 with a length of life with HD 1-317 months. The highest proportion of patients with longevity with HD 6-12 months.

Every time a patient undergoes HD, a CKF patient will be examined before HD (Per HD) and after HD (Post HD). Before HD the patient will be weighed to determine the increase in body weight from the dry weight which will be followed by checking the temperature, blood pressure and oxygen saturation level and pulse rate. During the HD process, the HD machine will set the amount of fluid withdrawal based on weight gain and the speed of blood

* Corresponding author.
E-mail address: teguh@unpkediri.ac.id

circulation (Quick of Blood). After the HD process is complete, the patient will be checked for blood pressure and weighed. Checking hemoglobin (HB) is also carried out at certain frequencies. If the patient's HB is below the minimum value, the patient will be advised to do a blood transfusion. So the inspection parameters to determine the quality after the implementation of HD.

The results of these examinations are recorded or recorded using a conventional approach. This means that the recording is done manually and kept at the health agency. So, CKF patients experience problems to find out the results of these examinations in real time. In addition, it is necessary to consult with a doctor to understand the results of the examination. Consultations with doctors also require scheduling as well so patients must also take the time to deal with this. In this study, a monitoring system is proposed that attempts to record the results of the inspection. This monitoring system is an embodiment of a medical record management information system that processes medical history data (Nurajijah & Riana, 2019). In addition, this monitoring system provides visualization of the data entered in the form of graphics. This monitoring system can also classify inspection results in the form of good or bad indications. So, this monitoring system tries to give CKF patients an overview of the results of their HD therapy and predict whether it is good or not good.

This study uses the Support Vector Machine (SVM) method to classify the results of the inspection after the implementation of HD. This classification method was chosen because it has a high level of accuracy for datasets with 2 classes (Nurajijah & Riana, 2019; Purnama et al., 2020). In this study using a linear kernel for the use of the SVM method. So, this monitoring system applies artificial intelligence with a machine learning approach. This has the goal that the knowledge base regarding determining the quality of examination does not need to transfer knowledge from experts (Ulfah & Anam, 2020). This is sufficient for a monitoring system to learn from the examination results dataset provided by the patient through input to this monitoring system. Therefore, changes in knowledge rules can be accommodated by the monitoring system without the involvement of experts, provided that the inspection parameters do not change.

This research is expected to contribute both conceptually and in practice. It is hoped that the results of this study can provide knowledge regarding the design of a monitoring system for CKF patients. It is hoped that this design will provide a pilot for building a more complex monitoring system for CKF patients so that they can find out real-time examination results and understand the visualization of existing data without consulting a doctor directly. In addition, this research is also expected to contribute knowledge regarding the stages of implementing SVM in the monitoring system and the results of evaluating the accuracy of the SVM method applied in this monitoring system.

2. Literature review

One of the learning methods made possible by the development of this transformation technology is smart learning, which aims to improve the quality of learning by using smart technology according to the learning context. This study focuses on a systematic literature review conducted as a pilot study in intelligent learning research for executives (Al Faruqi, 2019). The journals used in this review are various international journals obtained from several reputable journal databases. After going through a screening process with several inclusion and exclusion criteria, 15 journals were analyzed based on research questions. From this review it can be concluded that smart learning can improve the learning process in terms of motivation, engagement, and learning performance. Various information technologies are used to increase the efficiency and effectiveness of smart learning, which provides the learning needs of executive education.

The same research on kidney disease can also use an expert system to make a diagnosis. The Certainty Factor method can be used to overcome confusion in diagnosing patients (Fathushahib & Marselia, 2018). This uncertainty is a case study in the form of probability. Therefore it is necessary to have a combined certainty factor to treat the symptoms of the same disease. In addition, the Forward Chaining method was also applied in this research. This method is to describe the rules for diagnosing symptoms that are mapped from previous expert explanations.

The application of machine learning in the form of predictions is also applied to predict the location of broken rails on trains. In this study, the CRISP-DM method and several modeling techniques were applied which developed each other's accuracy results in detail. Then, classifiers used such as SVM (RBf kernel), Decision Tree, and K-Nearest Neighbor (Nafisah Nurul Hakim, 2020). Furthermore, Ensemble Learning was also applied in this study by selecting the Random Forest method. Evaluation results show that the use of the SVM method has the highest accuracy.

3. Research Method

This study uses an experimental approach. Experimental research is a research conducted to find the result of something that is done intentionally by researchers. Experimental research is also a research carried out to get the result of something that is carried out intentionally by researchers. The causes of all symptoms will be tested to find out the causes or independent variables that will affect the effects or the dependent variable (Thomas, 2021).

Figure 1 shows the stages of this research being carried out. This research begins with a literature study. It seeks to find references with the field of health monitoring systems and classification. The main reference sources used are papers published nationally. Next, comparing several previous studies with this study. There are several similarities and differences from previous studies with current research. In addition, there are several things taken from several previous studies carried out in this study.

The next step is data collection. Data from the implementation of HD will be inputted by health staff into the monitoring system. The collected data is stored in the database and data analysis will be carried out. The input data includes data before and after HD. Furthermore, the input parameters are body weight, blood pressure, complaints, body temperature, oxygen saturation level, heart rate, blood flow rate, target, hemoglobin content, and the number of transfusion bags.

The data analysis stage includes data cleaning and feature extraction processes. This data cleaning process is a preparation before heading to feature extraction. The raw data that has been input needs to be checked whether there are data whose attributes are still empty or the units are inconsistent. Clean data then needs to be divided into several attributes which are used as features or characteristics for the classification process. Data whose features have been defined also need to be labeled so that the system can study it to form a model.

The stages of applying the classification in detail are shown in Figure 2. The application of this classification is an implementation of the classification algorithm in the monitoring system development program code. So the application of this classification is also part of the process of building a monitoring system. The development of this monitoring system uses an Agile development approach so that it is more flexible in dealing with changes and can be worked on with the development team. The results of the development and use of this classification algorithm need to be evaluated to measure the performance of the monitoring system and the level of accuracy of the use of the classification algorithm. All of these research activities need to be documented and written down in a scientific article so as to contribute knowledge to future researchers. This scientific article will be published in a national level journal.

Figure 1 shows the classification flow in detail. The stages of data cleaning, data transformation, and data labeling are part of the data analysis process as shown in Figure 1. Data cleaning is done by uniforming units, naming conventions, filling in missing data, and adjusting sizes. Furthermore, the clean data will be transformed into existing numeric data and labeled. This labeling attempts to assign class values to data rows where this is useful for learning the system. The system will learn patterns from the data provided.

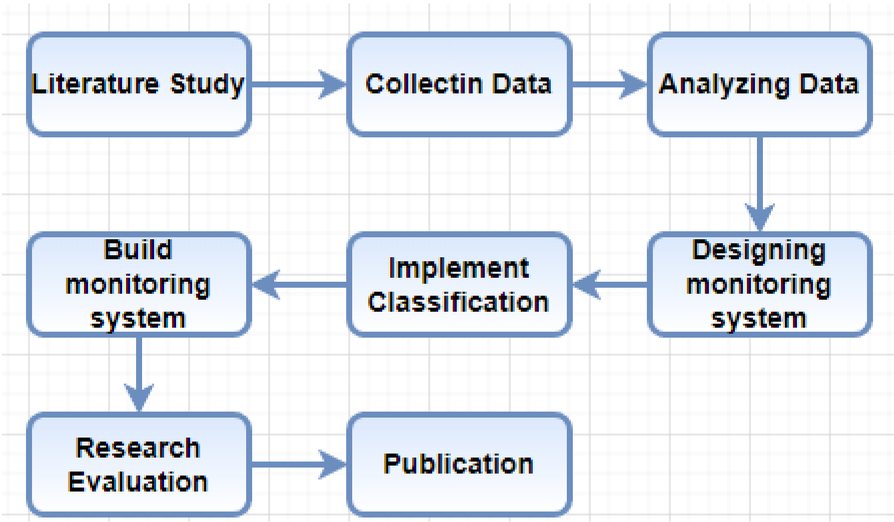

Fig. 1. Research Flow

Figure 1 shows the stages of this research being carried out. This research begins with a literature study. It seeks to find references with the field of health monitoring systems and classification. The main reference sources used are papers published nationally. Next, comparing several previous studies with this study. There are several similarities and differences from previous studies with current research. In addition, there are several things taken from several previous studies carried out in this study.

The next step is data collection. Data from the implementation of HD will be inputted by health staff into the monitoring system. The collected data is stored in the database and data analysis will be carried out. The input data includes data before and after HD. Furthermore, the input parameters are body weight, blood pressure, complaints, body temperature, oxygen saturation level, heart rate, blood flow rate, target, hemoglobin content, and the number of transfusion bags.

The data analysis stage includes data cleaning and feature extraction processes. This data cleaning process is a preparation before heading to feature extraction. The raw data that has been input needs to be checked whether there are data whose attributes are still empty or the units are inconsistent. Clean data then needs to be divided into several attributes which are used as features or characteristics for the classification process. Data whose features have been defined also need to be labeled so that the system can study it to form a model.

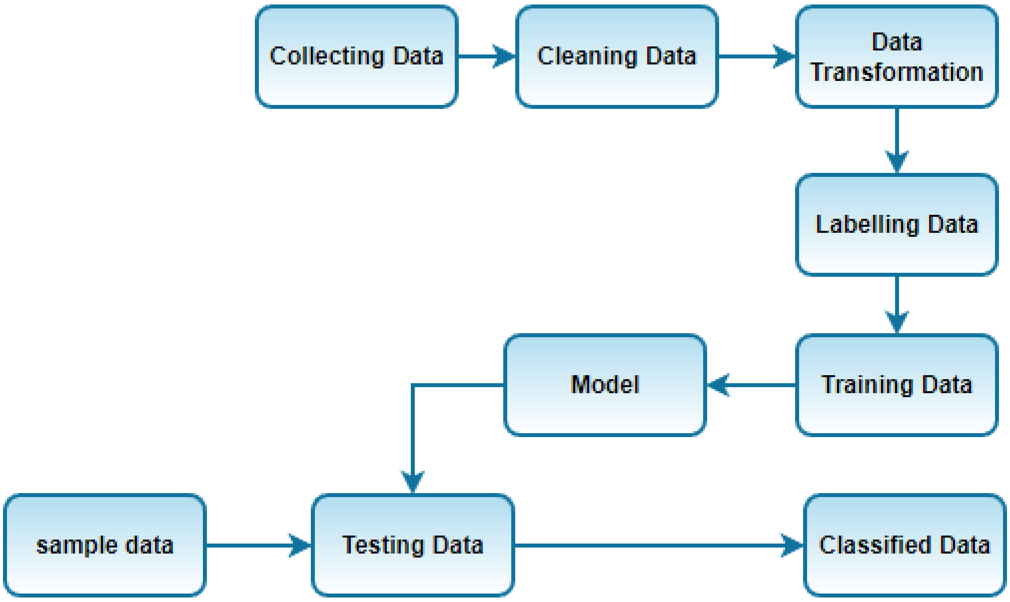

Fig. 2. Classification Flow.

The stages of applying the classification in detail are shown in Figure 2. The application of this classification is an implementation of the classification algorithm in the monitoring system development program code. So the application of this classification is also part of the process of building a monitoring system. The development of this monitoring system uses an Agile development approach so that it is more flexible in dealing with changes and can be worked on with the development team. The results of the development and use of this classification algorithm need to be evaluated to measure the performance of the monitoring system and the level of accuracy of the use of the classification algorithm. All of these research activities need to be documented and written down in a scientific article so as to contribute knowledge to future researchers. This scientific article will be published in a national level journal.

Figure 3 shows the classification flow in detail. The stages of data cleaning, data transformation, and data labeling are part of the data analysis process as shown in Figure 1. Data cleaning is done by uniforming units, naming conventions, filling in missing data, and adjusting sizes. Furthermore, the clean data will be transformed into existing numeric data and labeled. This labeling attempts to assign class values to data rows where this is useful for learning the system. The system will learn patterns from the data provided.

The classification process consists of data training, model generation, and data testing. Data training is carried out on data sets that have defined features and have labels. Then, the classifier will train the data along with the features and classes to generate knowledge. This is used as a monitoring system to determine the quality of HD performed by patients. This knowledge is at the core of monitoring system intelligence to determine outcomes. Knowledge representation if mapped in Figure 2 is a model. This model is used to classify test data that does not have a label. The classifier will label the test data without repeating the process from the beginning. So, the classifier can predict labels from previous raw data without using previous data sets or from previous physician rules.

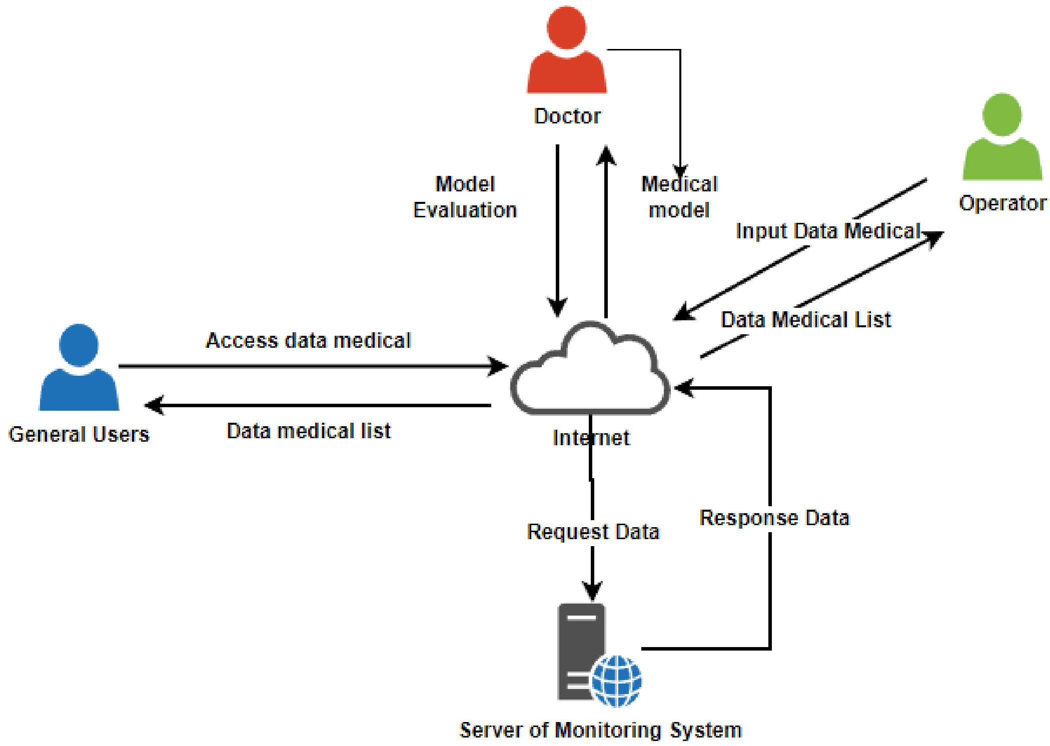

Fig. 3. Architecture of monitoring CKF Pation System

Figure 3 shows the architecture of the CKF patient monitoring system. It consists of 3 types of users namely CKF patients, Doctors, and Operators. Doctors can also monitor data and its visualization from CKF patients. Then, the operator here is a health staff whose job is to input data on examination results from CKF patients after carrying out HD therapy. CKF patients themselves have a dashboard that contains data before and after HD therapy with data visualization in the form and is accompanied by the quality of the therapy results labeled as good or not good.

This CKF Patient Monitoring System is in the form of a website that is placed on an online web server. So, access to the system is online, making it easier for all types of users to access it in various places, provided that they are connected to the internet. The system can also be accessed using a smartphone device of any specification via an existing browser. This is because the responsive display of the monitoring system adapts to the smartphone screen so users don't need to install it, just access the URL address in the provided browser.

In this study using the Support Vector Machine (SVM) algorithm as a classifier of the training data and test data provided. SVM is an algorithm that works using nonlinear mapping to transform the original training data into higher dimensions. In this new dimension, the method seeks a linear optimal separator hyperplane that is, “a decision boundary that separates data sets from one class to another. With a suitable nonlinear mapping to a high enough dimension, data from two classes can always be separated by a hyperplane. SVM finds this hyperplane using a support vector which is a set of important training output data and a margin that can be defined by the support vector.

The SVM algorithm completes the classification by finding the maximum marginal hyperplane [17]. Figure 4 shows two possible dividing hyperplanes and their associated margins. Before going into the margin definition, it is necessary to observe Figure 4 intuitively. Both hyperplanes can correctly classify any given data set. Intuitively, one would expect hyperplanes with larger margins to be more accurate in classifying the data set forward than hyperplanes with smaller margins. This is why during the learning or training phase SVM looks for the hyperplane with the largest margin, namely the maximum marginal hyperplane (MMH). Associated margins provide the greatest separation between classes.

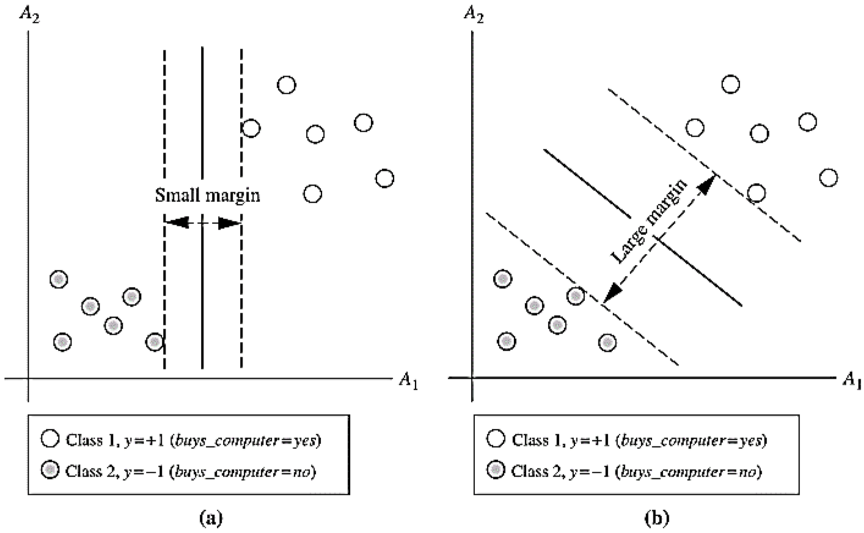

Fig. 4. Ilustration of hyperplane

Going to the informal definition of margin, it can be said that the shortest distance from the hyperplane to one side of its margin is equal to the shortest distance from the hyperplane to the other side of its margin, where the "side" of the margin is parallel to the hyperplane [17]. When dealing with MMH, this distance is actually the shortest distance from MMH to the closest training data of the two classes. A dividing hyperplane can be written by equation 1.

$$W.X + b = 0 \quad (1)$$

where W is the weight vector, that is, W = {w1, w2,..., wn}; n is the number of attributes; and b is a scalar, often referred to as a bias. To help with visualization, it is necessary to pay attention to two input attributes, A1 and A2, as shown in Figure 4 section (b). The training data is 2-D (eg, X = (x1, x2)), where x1 and x2 are the values of attributes A1 and A2, respectively, for X. If we consider b to be additional weights, this can be written as equation 2.

$$w_0 + w_1x_1 + w_2x_2 = 0 \quad (2)$$

So, every point that lies above the dividing hyperplane is fulfilled, this can be written in equation 3.

$$w_0 + w_1x_1 + w_2x_2 > 0 \quad (3)$$

Likewise, every point that lies under the dividing hyperplane is fulfilled, this can be written in equation 4.

$$w_0 + w_1x_1 + w_2x_2 < 0 \quad (4)$$

The weights can be adjusted so that the hyperplane that defines the side margins can be written as in equations 5 and 6.

4. Result and Discussion

Figure 5 shows the dashboard display of the CKF Patient Monitoring System. This dashboard has 4 main functionalities, namely patient profile, Medical Data, Data Analysis, and Consultation. The patient profile feature contains patient biodata consisting of name, address, place and date of birth, email, telephone number, dry weight, username and password. This can be done by the patient's own data processing. Furthermore, the Consultation feature contains fields to provide complaints where the complaint data will be received by the doctor.

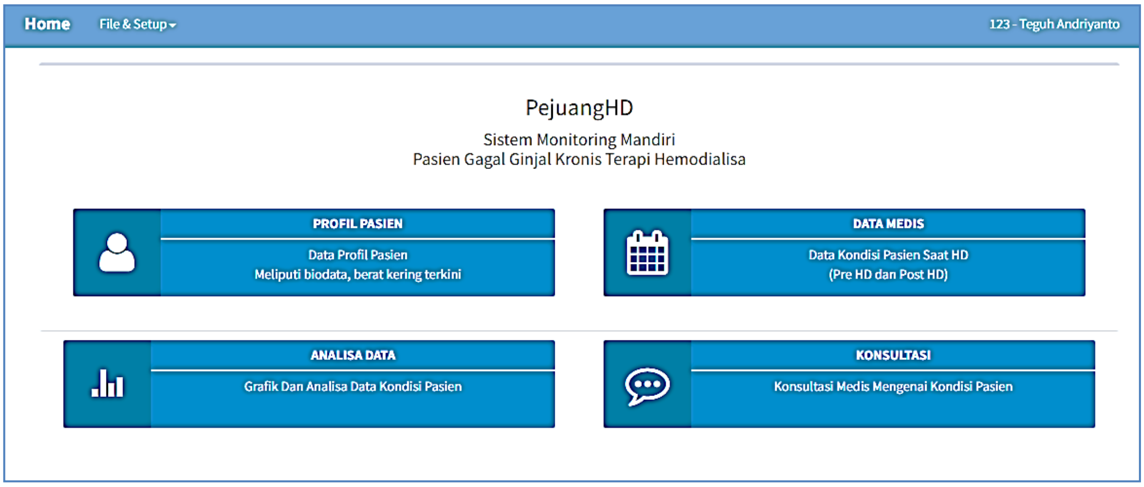

Fig. 5. Dashboard of CDK Pationt Monitoring System

The Medical Data Feature contains 2 sub features, namely Medical Data Input and View Medical Data. This is shown in Figure 5 part a. The medical data input form is shown in part b. Medical record staff or operators can enter data before and after HD therapy on this form. Historical data input is shown in Figure 5, section c. The data set in part c will be used as input for data analysis.

Figure 7 shows the display of the line graph of medical data that has been recorded. This is a visualization of medical data where there are 3 parameters that are recorded, namely red for orange for weight before HD. Blue for weight after HD therapy and green for dry weight. The data shown in Figure 7 indicates that the patient's dry weight has stabilized from the examination in the last 6 months.

The weight before HD therapy fluctuated but was always higher than the weight after HD and dry weight. For patients, it is expected that the weight before and after is always stable. This visualization contributes to CRF patients in knowing weight fluctuations before and after HD therapy.

Table 1 shows the results of the Confusion Matrix. This is a means to evaluate the type of classifier used. In Table 1 it is shown that the training data with labels is not good if it is classified using the model that has been formed it turns out to produce 52 data labels that match. However, there are 8 data generated from training data with poor labels that fall into the good classification. On the other hand, training data with the label Good is classified, resulting in 42 data labels that also match.

Table 1. Confusion Matrix

	NG	G
KB	52	3
B	0	42

Note:

NG : Not Good
G : Good

Table 2 shows the evaluation results using Cross Validation. In this evaluation using 3 fold. Then, the parameters used for this evaluation are precision, recall, F1 score, accuracy, and error. This evaluation produces a high level of accuracy, namely 96.91% with an error value of 3.09%.

Table 2. Evaluation Result using K-Fold Cross Validation

Category	Scores
Precision	1
Recall	0.945
F1 Score	0.972
Accuracy	96.91%
Error	3.09%

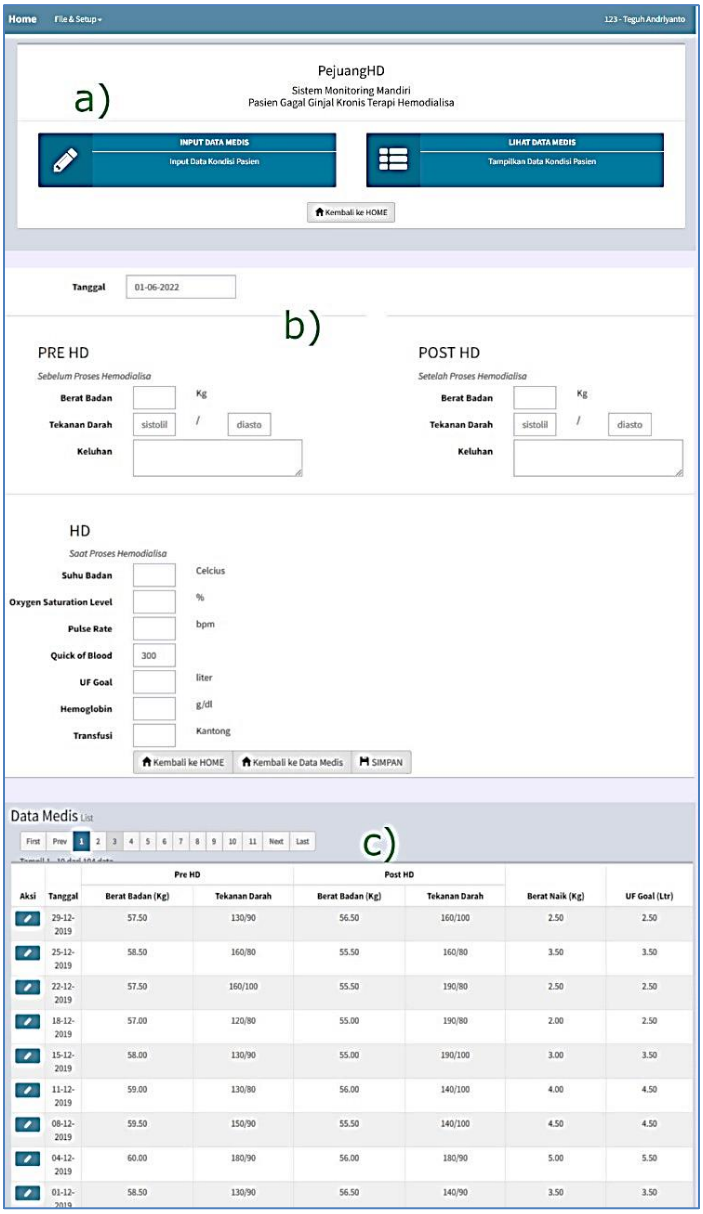

Fig. 6. Sub Features on Data Healt

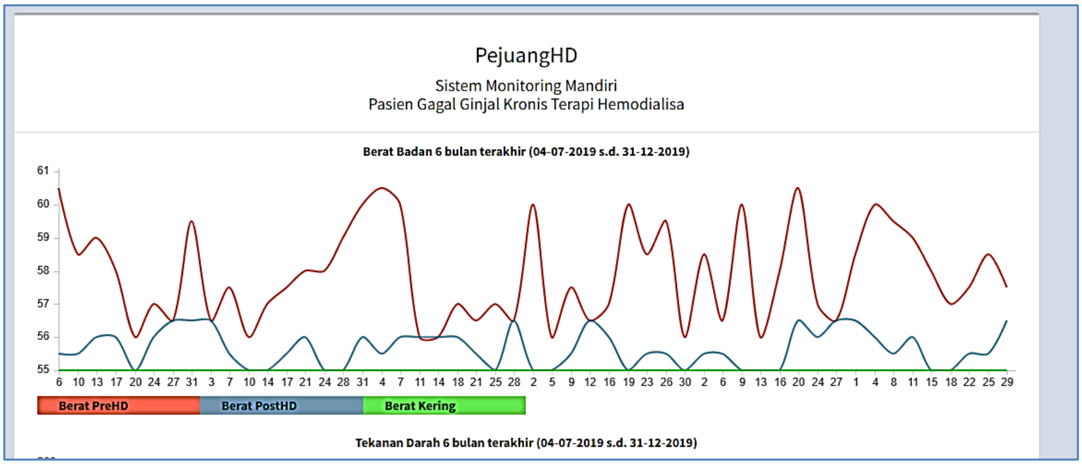

Fig. 7. Visualization Data Health

5. Conclusion

The CDK patient monitoring system has data analysis functionality which contains graphs which are visualizations of medical data that have been recorded. Patients can see directly the results of the visualization and more quickly understand the meaning of the graph because it has an explanation of the type of line displayed. In addition, the results of the evaluation of the use of the SVM algorithm as a classifier produce an accuracy level that can be used as a reference, namely 96.91%. This shows that the SVM algorithm can be used in a monitoring system for CDK patients because its accuracy is almost the same as the previous rules. If there is a change in the calculation rules for determining good or poor quality after HD therapy, then the system with the SVM algorithm can learn the given rule pattern provided that the parameters used as input in the calculation do not change.

There needs to be an improvement in the appearance of the CDK Patient Monitoring System, especially in the visualization of the data. In addition, there needs to be an evaluation of the existing model at a certain period so that the decrease in accuracy can be accommodated and the model can be updated periodically.

Acknowledgements

We would like to thank all members of the Nusantara PGRI Kediri University, especially the Research and Community Service Institute for supporting this research. In addition, we also thank the Jinav (Journal of Information and Visualization) Editorial Team for providing publication media for our scientific articles so that they can be accessed publicly.

References

Al Faruqi, U. (2019). Smart Learning Platform for Management Executive: A Systematic Literature Review. *Jurnal Sistem Cerdas*, *2*(3), 194–202. https://doi.org/10.37396/jsc.v2i3.46

Anonim. (2017). *Ginjal Kronis*. Direktorat P2PTM Kementrian Kesehatan RI. http://p2ptm.kemkes.go.id/kegiatan-p2ptm/subdit-penyakit-jantung-dan-pembuluh-darah/ginjal-kronis

Fathushahib, & Marselia, M. (2018). Perancangan Sistem Pakar Untuk Diagnosis Penyakit Ginjal Dengan Metode Certainty Factor Dan Forward Chaining. *Jurnal Sistem Cerdas*, *01*(02), 37–46.

Nafisah Nurul Hakim. (2020). Implementasi Machine Learning pada Sistem Prediksi Kejadian dan Lokasi Patah Rel Kereta Api di Indonesia. *Jurnal Sistem Cerdas*, *3*(1), 25–35. https://doi.org/10.37396/jsc.v3i1.58

Nurajijah, N., & Riana, D. (2019). Algoritma Naïve Bayes, Decision Tree, dan SVM untuk Klasifikasi Persetujuan Pembiayaan Nasabah Koperasi Syariah. *Jurnal Teknologi Dan Sistem Komputer*, *7*(2), 77–82.

https://doi.org/10.14710/jtsiskom.7.2.2019.77-82

Purnama, J. J., Nawawi, H. M., Rosyida, S., Ridwansyah, & Risandar. (2020). Klasifikasi Mahasiswa Her Berbasis Algortima Svm Dan Decision Classification of Her Students Based on Svm Algorithm and. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, *7*(6), 1253–1260. https://doi.org/10.25126/jtiik.202073080

Thomas, C. G. (2021). Research Methodology and Scientific Writing. In *Research Methodology and Scientific Writing* (Second). Springer. https://doi.org/10.1007/978-3-030-64865-7

Triaji, Y. W., Kridalukmana, R., & Widianto, E. D. (2017). Pembuatan Sistem Informasi Manajemen Klinik dengan Rekam Medis: Studi Kasus di Klinik Kebon Arum Boyolali. *Jurnal Teknologi Dan Sistem Komputer*, *5*(1), 15. https://doi.org/10.14710/jtsiskom.5.1.2017.15-22

Ulfah, A. N., & Anam, M. K. (2020). Analisis Sentimen Hate Speech Pada Portal Berita Online Menggunakan Support Vector Machine (SVM). *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, *7*(1), 1–10. https://doi.org/10.35957/jatisi.v7i1.196