Jurnal Mantik
Volume 4, Number 1, May 2020, pp. 494-504 E-ISSN 2685-4236
https://iocscience.org/ejournal/index.php/mantik

Analysis System About Muscle-Bone-Joint Symptoms by the Association Rule Method

Rian Harianto[1]*, Ucuk Darusalam[2], Andrianingsih[3]

Sistem Informasi,
Universitas Nasional, Jl. Sawo Manila, Kota Jakarta Selatan, Daerah Khusus Ibukota Jakarta 12520

Email :ryanzensai@gmail.com*, Ucuk_darusalam@citivas.unas.ac.id, andrianingsih@citivas.unas.ac.id

ARTICLE INFO

Article history:
Received: 04/05/2020
Revised: 20/05/2020
Accepted: 30/05/2020

Keywords:
Data mining,
Association Rule and
Algoritma Apriori

ABSTRACT

Muscle, Bone, Joint Problems are health problems in the human body. Problems of muscles, bones, joints felt by almost everyone, both parents, adolescents, and children. Symptoms of muscle, bone, joint problems can be pain in the knee, pain in the shoulder, pain in the neck, pain in the hands, pain in the legs or pain in the back. To categorize the problems of muscles, bones, joints, there are many problems in management reporting to find out the symptoms that often occur in one semester. Based on this it is necessary studies that lead to the creation of a system that can categorize the spread of symptoms in muscles, bones, joints to determine the results of diagnosis based on the pattern of symptoms experienced by patients. So that helps management to find out and get information based on the pattern of symptoms that occur in patients. One alternative to the problem, this study we make a search pattern or association rule association (associative rules) of large-scale data and is very closely related to data mining that can be used to find certain rules that associate data with one other data with a priori algorithm method. The purpose of the method is to search historical data to identify data patterns based on previous characteristics to determine the disease diagnosis. So that the information generated can be used by management to provide appropriate diagnostic information according to the natural pattern of symptoms by the patient.

1. Introduction

The problem with symptoms in muscles, bones, joints in the human body, is one of the symptoms that make the body feel pain. Symptoms that arise due to problems of muscles, bones, joints also vary, from pain in the limbs, such as pain in the knees, hands, waist, neck, legs, or shoulders. Based on this, there are problems in the medical reporting system in knowing the spread of symptoms of the disease which often occurs in one semester on medical data reports in Flex-Free clinics. Reporting that is done at this time is still by manual input by the medical team. So it takes time to process it because of the large amount of data obtained. Because of the problems that often occur today. The management team plans to build a system so that it can help the medical team not to input manually enter data manually. So to help the medical team in knowing the results of reports of medical symptoms that occur, we need a system to be able to help and lead a clear and accurate analysis to determine the diagnosis results of the disease. So that it is expected to be able to help the medical in getting disease diagnosis information based on the pattern of symptoms experienced by the patient and also help the performance of the medical team in providing more accurate prevention and treatment. As an alternative way to the solution of this research is to make a pattern or relationship Rule (associative rules) of large-scale data associated with data mining. So it is expected to find some rules to be able to categorize the data one with the other data with an algorithmic method. In its application is to use a priori algorithm, is expected to be able to search historical data, identify diagnoses of the disease based on the pattern of symptoms suffered by the patient. Then the results are expected to get the right information and can be justified. So that information can help the medical team in conveying medical information to patients.

2. Theoretical Basis

2.1 Understanding Data Mining

In this study, we tried to conduct research with the data mining association rule method, which is a method of managing information from a large database that includes: data extraction processes, data recognition, data analysis, and presentation of information that can help decision making urgent. In general, it can also be interpreted that the data mining method is also called the Knowledge Discovery Database (KDD)

or it can be called a series of processes used to extract important and interesting data patterns from large amounts of data obtained to obtain knowledge that had not been known at the time it was carried out. Manually. The use of data mining methods is used as a solution after many problems from data owners who have accumulated data that has been collected for years, for example patient data, patient medical record data, employee data or sales transaction data. After the accumulation of data, the idea of what data should be done to the data and how to use the data in order to help data processing more quickly and efficiently in the reporting process, as well as help get accurate and easy to understand information.

2.2 Steps in Implementing Data Mining

a. Selection Data
   Stage 1 of data mining is the data selection in which this step is, some suitable data is analyzed and then the database is taken. Then correlation analysis is performed to find new features. So that the data attribute is checked whether the data attribute is redundant or will not be used. So that the attribute of the data is independent, which means that the attributes of one another do not influence or are interconnected.
b. Data Preparation
   In the second stage, data extraction is carried out. This stage is carried out to extract the data and transform it into a data mining algorithm. One of them is a joint table, adding new fields and cleaning up unused data.
c. Data Integration (Combining Data)
   In the process of this step to combine several data from several sources. After that the data to be combined is stored in a coherent data storage. Macem in data storage is included in Multiple Databases, Data Cube or Flat Files. These are done for data integration with object matching schemes and data redundancy.
d. Data Transformation
   Data transformation is a step, in which data will be transformed or consolidated into the formation of appropriate and appropriate data to be chosen that can be used (mined).
e. Data Mining
   Data mining is a process of important data, which at this stage the data method is applied and applied appropriately to find a data pattern.
f. Pattern Evaluation (Pulling data to gain knowledge)
   At this stage, an interesting data is identified to be applied or get a knowledge based on some suitable measurement analysis.
g. Pattern Deployment (developing data patterns)
   In the Pattern Deployment step, the analyzed data is developed to find new data patterns. The next one is designed, so that it can be used to determine the Decision Support System (DSS). So that it can be made a report and knowledge with the aim to do further processing.
h. Pattern Monittoring (Monitors data conclusions)
   At this stage the main conclusions obtained from the results of the development of data mining. Where similar patterns obtained can be applied to conditions that occur in the future.
i. Knowledge Presentation
   In this final stage, visualization techniques are performed to present the knowledge of the results of data mining analysis in the form of applications to users.

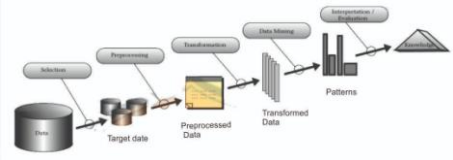


Fig 1. Steps in the Data Mining Process in Bantuk Diagram

3. Research Methods

3.1 Data collection

The data collection was carried out from the results of the 2019 medical census data at the Flex-Free Musculoskeletal Rehabilitation Clinic, which is a special rehabilitation clinic that deals with muscle-bone-joint problems. In the next process the analysis process is carried out by following a number of techniques such as the following:

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

a. Association Rule

In the Analysis this time the application of Data Mining is conducted using the Association Rule methodwith the aim of getting an information from each items that are interconnected into the form of rules. Thus it is expected that the use of the Association Association method can find the results of diagnosis based on the pattern of data symptoms that occur in patients. . So it takes a long time and work in determining the diagnosis results based on data in the symptom pattern. So to determine the relationship an efficient algorithm is made.

Table 1
Association parameters in the category of symptoms

Disease	The symptoms		
Nerve Clamps (HNP)	1.	Back pain	G1
	2.	Neck pain spreads from the neck to the fingers	
	3.	Rada tingling and numbness arises	
	4.	The muscles in the neck feel stiff and weak	
	5.	Headache	
	6.	Urinary disorders or bowel movements	
Osteoarthritis (OA)	1.	Swelling in the joints	G2
Mixing	2.	Muscle weakness and reduced muscle mass	
	3.	Joints feel stiff when I wake up in the morning or after a long period of movement	
	4.	Pain after the joint is moved (pain usually occurs in the knee and hip joints)	
	1.	Tungggung pain	G3
Osteoporosis	2.	Slouching posture	
	3.	It's easy to experience a broken bone, even if it's just a mild collision	
Uric acid	1.	Joints suddenly feel very sick	G4
	2.	The skin between the affected joints will appear, scaly, peeled and itchy	
	3.	The pain is accompanied by swelling and stiffness	
	4.	Sufferers will have difficulty walking due to disruptive pain, especially at night	
Scoliosis	1.	Body posture looks sideways to one side (S-shaped) when viewed from behind	G5
	2.	Impaired appearance (aesthetic)	
	3.	Low back pain	
	4.	Nerve disorders	
	5.	Impaired lung development	

Association Rule method is a technique in data mining to find associative rules with a combination of items. This associative rule is used with the aim of finding out several diagnoses based on the data pattern of symptoms experienced by the patient. In knowing an associative rule can only be done with two parameters, namely the Value of Support (supporting value), the value in the percentage combination of items in the database, while the Confidence Value (the value of certainty) is the value of the strength of the relationship between items in the associative rule process. So the Associative Rules are stated in the form:
Example :

{ Neck Pain and Back Pain}> {Nerve Clamps} (Support = 40% Confidence = 50%)

Which can be interpreted: 50% of the data in the database that contains items **Symptoms of Neck Pain and Back Pain,** also includes diagnosis **Nerve Clamps.**While Support 40% of all data in the database, contains the results of the diagnostic items. So it can be interpreted as well: "If a patient who experiences symptoms of neck pain and back pain is likely to be 50% diagnosed with Pinch Nerve. In this rule can be interpreted to represent 40% of data from medical records. So that the analysis of associative rules can be interpreted as a process to find all associative rules that meet the minimum requirements for support and minimum requirements for confidence.So that the basic associative rules are divided into 2 stages:

a. High Frequency Pattern Analysis

At this stage is done to determine the Item Combination that corresponds to the minimum value in the support value of the symptom data. The formula to get the value of Support can be done with the following formula:

Number of Symptoms to contain (A)

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

$$\textbf{Support (A)} = \frac{\textbf{Total Diagnosis of Disease}}{}$$

The support value of these 2 items is obtained from the following next:

$$\textbf{Support (A, B)} = \textbf{P (A} \square \square \textbf{ B)} = \frac{\textbf{Number of Symptoms it contains (A\&B)}}{\textbf{Total Diagnosis}}$$

b. Establishment of Associative Rules

After the first process in the high frequency pattern is carried out, the next step is to determine the Associative Rules to fulfill the minimum Confidence value. To calculate the value of confidence, then the Association (A&B) rules with the following formula:

$$\textbf{Confidance} = \textbf{P (B / A)} = \frac{\textbf{Number of Symptoms it contains (A\&B)}}{\textbf{TOTAL DIAGNOSIS CONTAINING (A)}}$$

In the formula above, it can be interpreted that the value containing the confidence can be obtained by using the method, a division of the amount of data containing the value of Item (A) and the value of Item (B) (Value of the Amount containing A&B) with the amount of data containing Item (A).

The next step is searching for the Apriori Algorithm. Apriori Algorithm is one algorithm that is commonly used to carry out the search processFrequent Itemset with Association Rules. Apriori Algorithm Method uses Level-Wise Search approach in the search process. Where K-Itemset is used to obtain (k + 1) Itemset. In this process it is carried out until no more combinations are formed. So that the formation of the association pattern with the Apriori Algorithm. There are 2 stages, namely looking for Frequent Itemset (the set that meets the minimum value of Support) for stage 1. Next step 2 is forming the associative pattern of the frequent Itemset that has been obtained with the Confidence value.

c. Application System Development

Application System Developmentis a stage in building a software or application in helping the data information needs to perform data processing. So that the system built, is a system device that can implement data with the Association Rule Mining method using the Algorithma Apriori calculation to find the diagnosis items based on symptoms. So it is expected to find the results of disease diagnosis information based on the analysis of perceived symptom patterns. In determining the analysis design models are made to develop the system and the stages include: Planning (planning), design (design), coding (coding) and testing phase (testing).

4. Analysis and Discussion

4.1 Rule Association Rule Analysis

In this discussion I am trying to retrieve symptoms diagnosis data at the 2019 medical data census at the Flex-Free clinic. The data has described some of the various causes of symptoms that plague the diagnosis of a disease. For this reason, in this analysis, I try to do an analysis by applying the Data Mining and Association Rule method, which aims to find information on interconnected items that form a Rule. So the Association Rule that is applied to the pattern of data based on symptoms will present a disease. So the process of finding the relationship between symptom items and disease needs repeated readings.

For this reason, in this study I used 5 data samples of diseases in which each disease has a pattern of symptoms in determining the disease:

Table 2
Disease Data Respentation

Patient	Disease Item
1	Nerve Clamps (HNP)
2	Osteoarthritis (OA) or Calcification
3	Osteoporosis
4	Uric acid
5	Scoliosis

If formed in a Tabular Data Table based on the patient distribution it would look like this

Table 3
Tabular data format based on patient distribution

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

Patient	Nerve Clamps (HNP)	Osteoarthritis (OA) or Calcification	Osteoporosis	Uric acid	Scoliosis
1	1	0	0	0	1
2	0	1	0	1	0
3	1	0	0	1	0
4	0	0	1	1	1
5	0	0	1	0	1
6	1	0	0	1	0
7	0	1	0	1	0
8	1	0	0	0	1
9	0	1	1	0	0
10	0	1	0	1	0
11	0	0	1	1	0
12	1	0	0	0	1
13	0	1	0	1	0
14	0	0	1	0	0
15	0	0	0	0	1
16	1	0	0	1	1
17	0	1	0	1	0

From the tabular data above, it can be interpreted "is the patient presented in table 2, where each patient is categorized based on the type of diagnosis, then the set of item A (diagnosis of nerve cramps) and set B symptoms one of which (back pain). Then the association rules will be formed:

"If Diagona has a nerve disease, then one of the symptoms includes back pain."

4.2 Application of Apriori Algorithm

In the application of a priori algorithm aims to obtain rules that have support and confidence values. In this analysis, the two rules have a minimum value of 20% support and 35% confidence. Then the itemset of the set of Items in I, and K-ItemSet only contains K-Items. Then the results are like the following table example:

Table 4
Candidate 2 Item Set

Combination	Count	Combination	Count
Nerve Clamps (HNP)	2	1. Back pain 2. Neck pain spreads from the neck to the fingers 3. Rada tingling and numbness arises 4. The muscles in the neck feel stiff and weak 5. Headache. Urination or bowel movements	2
Osteoarthritis (OA) or Calcification	3	1. Swelling in the joints 2. Muscle weakness and reduced muscle mass 3. Joints feel stiff when I wake up in the morning or after a long period of movement 4. Pain after the joint is moved (pain usually occurs in the knee and hip joints)	3
Osteoporosis	1	1. Tungggung pain 2. Slouching posture 3. It's easy to experience a broken bone, even if it's just a mild collision	1
Uric acid	2	1. Joints suddenly feel very sick 2. The skin between the affected joints will appear, scaly, peeled and itchy 3. The pain is accompanied by swelling and stiffness 4. Sufferers will have difficulty walking due to disruptive pain, especially at night	2
Scoliosis	1	1. Body posture looks sideways to one side (S-shaped) when viewed from behind 2. Impaired appearance (aesthetic) 3. Low back pain 4. Nerve disorders 5. Impaired lung development	1

From the results of the data table above, if determined with a value of P = 3 then in F2 = includes: {{Nerve Clamps} {Osteoarthritis}, {Osteoporosis} {Gout} {Scoliosis} Following the table:

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

Table 6
Prospective 3-Item set

Combination	Count
Nerve Clamps (HNP)	2
Osteoarthritis (OA)	3
Osteoporosis	1
Uric acid	2
Scoliosis	1

4.3. Establishment of Associative Rules

The next stage is the formation of associative rules, after all high frequency patterns have been discovered. Then look for associative rules that meet the minimum requirements for confidence by calculating the confidence of associative A&B rules with a formula like the following:

$$\textbf{Confidance} = P(B/A) = \frac{\textbf{Number of Symptoms containing A\&B}}{\textbf{Number of Diagnoses Containing A}}$$

From the results of the formula above, we get the results of the amount of support and confidence from the association rules as shown in the following table:

IF Antecedent Then Confidence	Support	Confidence
If diagnosed with nerve pain then the symptoms are back pain and neck pain	2/11 = 22%	2/4 = 50%
If diagnosed with Osteoarthritis then Symptoms of Pain in the Knee and Swollen Joints	3/11 = 27.3%	3/4 = 75%
If diagnosed with Osteoporosis then Symptoms of Pain in the Knee and Bone Loss	2/11 = 22%	2/4 = 50%
If the diagnosis of Scoliosis then Symptoms appear oblique posture, low back pain and nerve disorders	1/11 = 11%	1/4 = 25%

From the table it can be interpreted that the minimum confidence value is 25% which refers to the diagnosis of osteoporosis and scoliosis, then this rule can be formed with 2 antecendents, namely:
"IF diagnosis of osteoporosis then symptoms of pain in the knee and broken bones"
"IF Diagnosis of Scoliosis then the symptoms are oblique posture, low back pain and nerve disorders"

4.4. New System Analysis

In developing the system this time I tried to develop a system based on Java NetBeans IDE 6.8 and MySQL database. In developing the muscle-bone-joint symptom analysis system this time, the flow of the process of the system applies the association rule mining method with a priori algorithm analysis. With the aim of obtaining the following results:

- This system is intended to be able to analyze the database of patient diagnoses according to the pattern of symptoms
- This system is intended to get the output of several items that are interrelated with the pattern of symptoms experienced by patients, so that it can refer to the results of disease diagnosis.
- This system is expected to produce information from the results of analysis between combinations with items in the form of patterns of symptoms suffered by patients, sso that you can find out the results of the analysis of the disease based on each disease category and the number of patients.

4.5. System Diagram and Design

a. Use Case Diagrams

In the system flow activity there is only one user, with activities including inputting patient data, analyzing patient diagnoses and viewing patient history reports. But from the system activities there is one activity that Depands On. What can be interpreted as these activities can be done after the previous activity is done.
Example diagram:

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

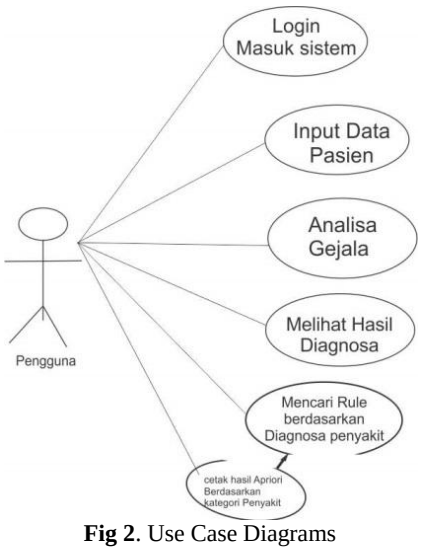


Fig 2. Use Case Diagrams

1) Use Case Definition

The Use Case diagram description is described in the following table:

No.	Use Case	Description
1	Login to the system	Here the user must enter a Username and Password In using the system
2	Patient data input	Users must input patient data, before analyzing the diagnosis of the disease
3	Symptom analysis	The user analyzes the symptoms to get an appropriate disease diagnosis
4	See the results of the diagnosis	Users here can see the results of patient diagnoses that have been inputted and generated from the database
5	Looking for rules based on the criteria for diagnosing the disease	Here the user searches for rules based on the results of diagnostic analysis. In this activity the user is presented with data information to find out the number of patients based on the results of the diagnosis, then the system will determine it.
6	Print a priori results	In this activity the User can pass the output of the a priori algorithm based on the rule search that has been done before. These results are obtained, if the activity of looking for the rule of disease symptoms has been carried out.

2) Activity diagram

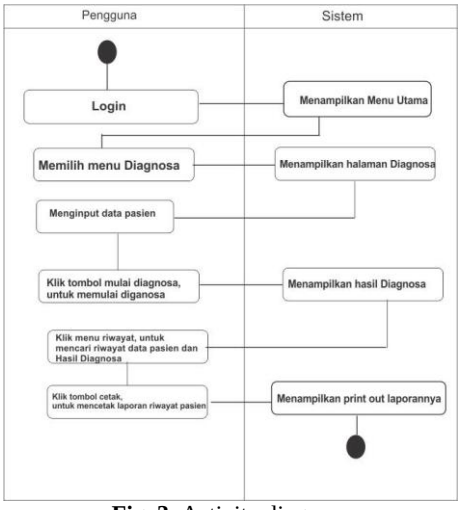


Fig 3. Activity diagram

3) Diagnosis Flow Chart Based on Symptoms

In the system flow obtained there is a system flow in the form of a flowchart whose contents illustrate the flow of the system with the association rule mining method. So it can be seen from the picture as follows:

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

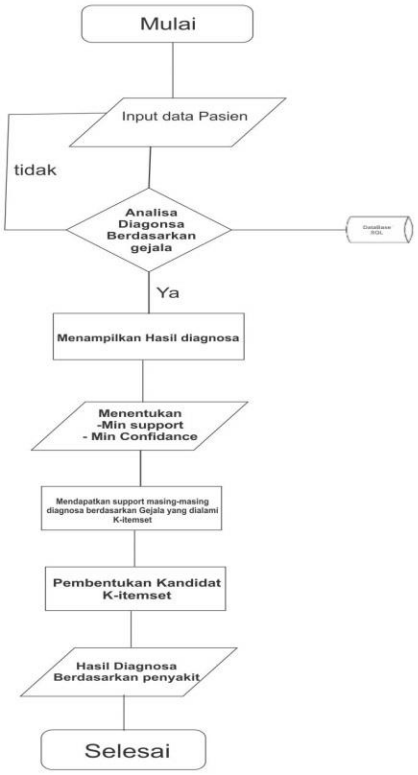


Fig 4. Flowchart Flow Travel System

4.6. Results

In this system design, its implementation is carried out with an interface data input model in which the application workflow process is in accordance with the data needs to be analyzed and the design has been carried out. As well as already applying the Assosiation Rule Mining method and a priori algorithm calculation. Here is the system display:

Fig 5. Example display Login

Login view is used to start and log in to the main system by using user and password authentication that has been registered in the user database.

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

Fig 6. Display the Main Menu

The main menu page is the start page for selecting menus, after logging in. The main menu page is the main system for conducting data analysis processes for using the Rule Mining Association.

Fig 7. Disease menu display based on the pattern of symptoms

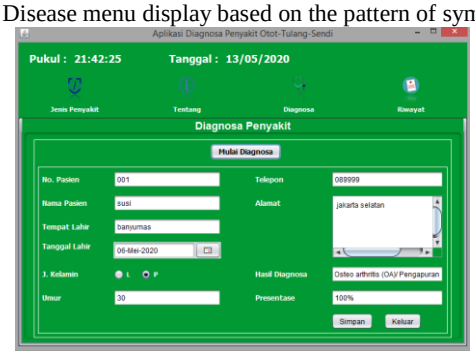


Fig 8. The menu display determines the diagnosis of disease

Used to analyze the diagnosis of the disease based on the pattern of symptoms felt by the patient

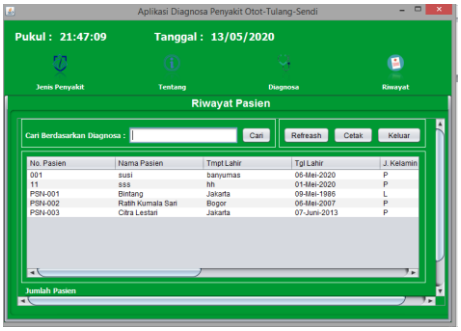


Fig 9. Display patient data history

This display is used to display patient history data and the results of diagnosis

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

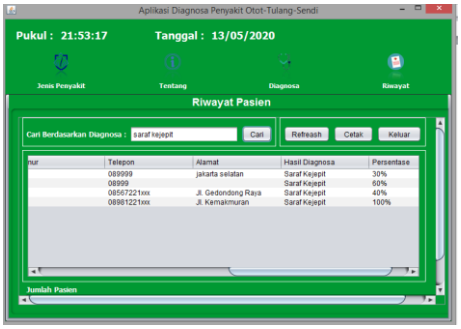
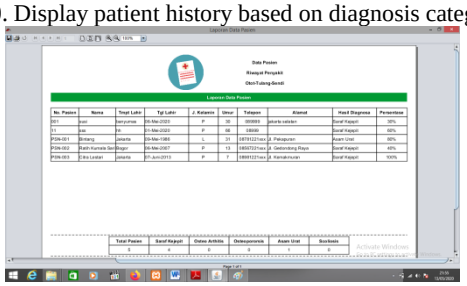


Fig 10. Display patient history based on diagnosis categories



Picture 10. Display Print Out report

5. Conclusion

From the results of the research and discussion that has been carried out on the analysis of the data pattern of Muscle-Bone-Joint Symptoms by applying the Association Rule method, so the analysis results are obtained as follows:

a. In applying the Association Rule data mining method, it can help analyze diagnoses based on the pattern of pain symptoms experienced by the patient.
b. The use of the Association Rule Method helps in estimating the diagnosis of the patient's disease.
c. In developing an application system with the Data Mining method it can help in knowing the number of diagnoses per disease category.
d. In applying the Association Rule method, it can be used to search for information in the future.

6. Reference

[1] Abdallah Alashqur, "MINING ASSOCIATION RULE: A DATABASE PERSPECTIVE", International Journal of Computer Science and Network Security, Vol 8 No. 12, December 2008

[2] D.Fitriati, "IMPLEMENTASI DATA MINING UNTUK MENENTUKAN KOMBINASI MEDIA PROMOSI BARANG BERDASARKAN PERILAKU PEMBELI DENGAN MENGGGUNAKAN ALGORITMA APRIORI", Annu Resarch Semin,vol. 2, no. Implementasi Data Mining, pp. 472–480, 2016.

[3] D.Listriani, A.H. Setyoningrum, and F.E.M.A, "PENERAPAN METODE ASOSIASI MENGGUNAKAN ALOGRITMA APRIORI PADA APLIKASI ANALISA POLA BELANJA KOSUMEN (STUDI KASUS TOKO BUKU GRAMEDIA BINTARO)", Tek. Inform., vol. 9, no. 2, pp. 120–127, 2016.

[4] Harris Kurniawan, Fujiati, Alfa Saleh " ANALISA POLA DATA PENYAKIT RUMAH SAKIT DENGAN MENERAPKAN METODE ASSOCIATION RULE MENGGUNAKAN ALGORITMA RULE", STMIK POTENSI UTAMA, SEMINAR NASIONAL INFORMATIKA 2014

[5] Kusrini, "PENERAPAN ALGORITMA APRIORI PADA DATA MINING UNTUK MENGELOMPOKKAN BARANG BERDASARKAN KECENDRUNGAN KEMUNCULAN BERSAMA DALAM SATU TRANSAKSI", YOGYAKARTA,2007

[6] O.S.A Destiyati and E.Ariwibowo, "ANALISA PERBANDINGAN ALGORITMA APRIORI DENGAN ALOGORITMA HASH BASED PADA MARKET BASKET ANALIS diAPOTIK UAD", Sar.Tek Inform, vol. 3, no. 1, pp. 1–10, 2015.

[7] Therling. K, " INTRODUCTION TO DATA MINING DISCOVERING HIDDEN VALUE IN YOUR DATA WAREHOUSE", www,therling.com, 2006

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.