DATA MINING APPLICATION FOR DISEASES CLASSIFYING BASED ON REGION USING CLASSIFICATION METHOD AND C4.5 ALGORITHM

Danang Aristianto[1], Alif Finandhita[2]

[1,2] Teknik Informatika – Universitas Komputer Indonesia
Jl. Dipatiukur 112-114 Bandung
E-mail : danang.aristianto@gmail.com[1], finandhita@gmail.com[2]

ABSTRACT

Puskesmas Purwakarta is one of Community-Healt Center in Purwakarta. There are many patients come from various regions in Purwakarta every days. The amount of them can be hundreads for each service in one day and gives many daily visit repots. Every month, Puskesmas purwakarta holds an extension about desease for the community in the region. But the information given is not what the community needs. The problem occured because the extension does not refer to the daily visit reports.

Data mining is the computational process to discover hidden information in large data sets. The methods used in this research is classification method. The writer uses this method to classify diseases based on region that is found from daily visit reports. The algorithm used in this research is C4.5 algorithm. This method and algoritmh will produce a result tree that can help Puskesmas Purwakarta to make a better extension based on region from the daily visit reports.

The research compares the result tree produced manually with the result tree produced by an application. According to the research outcome, the application can help Puskesmas Purwakarta to hold on extension based on region from daily visit reports that have been proceeded by the application.

Kata Kunci : C4.5 Algorithm, Classification, Data Mining

1. INTRODUCTION

1.1 Background of Study

Puskesmas Purwakarta is one of Community-Health Center in Purwakarta. It is located at Jalan Siliwangi no 4 Purwakarta. Just like common health center, Puskesmas Purwakarta gives services to the community every day. There are three kinds of services for the patients. The services are general service, Jampis (Jaminan Purwakarta Istimewa) and BPJS (national health care security).The administration system of three of them are different. The people should pay the administration to Puskesmas Purwakarta for general service, but Jampis and BPJS are free. Each day, Puskesmas Purwakarta takes care many patients from various regions around Purwakarta. The amount of them can be hundreds for each service in one day and gives many daily visit reports.

From the reports, Puskesmas Purwakarta needs further data-diseases tabulation to discover new information for making a decision about making extension in certain region according to data classification that have been constructed. Puskesmas Purwakarta usually held an extension based on their annual program. Few diseases will be informed in the extension. The disease are Diabetes, Hypertension, Pneumonia, Diarrhea and TBC (Tuberculosis).The previous fact shows that the information given in the extension was not what the community needed to know. That was caused by the information given was not based on the daily visit reports.

Puskesmas Purwakarta needs a processor for their daily visit reports. It will help them to classify the diseases for making decision which region they should hold an extension. The data processing used for this case is Data Mining and Classification method [1]. The use of classification method is used for classify diseases based on region from daily visit reports. Thus, the diseases information for the extension can reach the target due to what the community needs. There writer uses C4.5 algorithm because it is fit the classification method and decision tree construction [2] [3]

Using data mining with classification method and C4.5 algorithm may be able to be used by Puskesmas Purwakarta as a helpful tool for making right decision on holding an extension based on the information that community needs.

2. PAPER CONTENTS

2.2 Objective and Purpose

Based on the case, the objective of creating this application is to apply data mining for classifying diseases according to region with classification method and C4.5 algorithm.

The purpose of constructing this application is :

In order for disease information submitted by the Puskesmas Purwakarta to the community through

counseling activities is on target and help Puskesmas Purwakarta making a decision to put the information for their extension in certain region based on classification data from the data mining.

2.3 Data Mining

According to Daniel T Larose's 'Discovering Knowledge in Data: Introduction to Data Mining', data mining is a process to discover connections, patterns and preferences by checking large data sets in the database with statistics and mathematics [1].

Some early definitions of data mining including focus on the automation process. Data mining, in Berry and Linoff's Technique for Marketing, Sales, and Customer Support, is a exploration process and automatic or semi-automatic analysis to large data sets in order to discover an significant pattern or regulation.

2.4 Classification

In classification, based on Daniel T Larose's 'Discovering Knowledge in Data: An Introduction to Data Mining', there are category variable target or class target [1]. The example for classification is income arrangement can be divided into three categories, high income, medium income and low income. Data mining model checks a large set of data, every data contains target variable information or class target. For example, a researcher wants to get classification of income arrangement from someone who is not in the database based on characteristic of that person like age, sex, and job. This task fits the data mining and classification method. The algorithm process is more or less like this. First, checking the data that contain predicator attribute (have not been clarified) and the data have been clarified its income arrangement. With this way, algorithm (software) will 'learn' about attribute combination that related with the income arrangement. For example, an older woman maybe related with high-income class. This data known as training set. Then the algorithm will search new data without information or income arrangement. According to the classification in training set, algorithm will place classification result to the new data [1].

2.5 C4.5 Algorithm

Algorithm C.45, based on Daniel T Larose's 'Discovering Knowledge in Data : An Introduction to Data Mining', is development from algorithm ID3. Algorithm C4.5 and ID3 were created by J. Rose Quinlan on 1970s to make a decision tree. Algorithm C4.5 uses Entropy and Gain concept to make an optimal decision tree [1]. Algorithm C4.5 to construct a decision tree are :

- Choose attribute as a root
- Make a branch for each point
- Divide the cases into branches
- Repeat the process for each branch until all the cases have the same class

To calculate the entropy can be seen on this equation :

$$Entropy(S) = \sum_{i=1}^{n} - pi * log_2 pi$$

Information : S : cases association

n : partition amount S

pi : proportion from S_i to S

To calculate the gain use this equation :

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^{n} \frac{Si}{S} * Entropy(Si)$$

Information : S : cases association

A : attribute

n : attribute patition amount A

S_i : cases amount in partition – i

S : cases amount in S

2.6 Decision Tree

Decision Tree are the transformation of the data. The construction starts from the root node that is placed on the top of a decision tree and it will construct a branch. Each branch contains an attribute point and will aim to the other node or directly aim to the leaf node or class [1]. The material for constructing the decision tree is a table with attribute and record. Each attribute on the table explain a decision class.

2.7 Data Mining Steps

The writer uses KDD (Knowledge Discovery in Database) on this research. KDD is a field to integrate various methods that involves statistics, data base, artificial intelligent, machine learning, pattern recognition, modeling handling-uncertainty, data visualization, optimization, management information system and knowledge based system [5].

2.7.1 Data Selection

Data selection is an assortment from operational data sets that is compatible for data mining requirement. On this research, the data chosen are daily visit reports, which represent almost all of the data, at Puskesmas Purwakarta. The daily visit reports used is taken from the data on July 2015 and the sample is the report on July 31 2015.

2.7.2 Preprocessing

This process aims to select one attribute from a data and removes duplicative data also to fix an error on the data such as unfilled attribute or typo. The table below explains attribute choosing from daily visit reports at Puskesmas Purwakarta.

Table 2.1. Attribute choosing table

Name of Attribute	Explanation
Nama Penyakit / Disease name	Disease name attribute explains the types of disease that is suffered by the community.
Umur / Age	This attribute is chosen because it can explain how to differ the

	treatment for the patients such as medicine giving to underage and aduvlt that have different dose.
Wilayah / Region	Region attribute is the decision determinant for the extension that shows a kind of disease around Puskesmas Purwakarta (Nagri Kidul, Nagri Tengah, Sindangkasih, and Luar Wilayah).

2.7.3 Transformation

Transformation is a process to change a value in an attribute with new suitable value for data mining process. For example, age attribute does not change because the value of this attribute already inside the age interval. Therefore, it is already fit the data mining process. Region attribute has also already divided patient's addresses into each region according to Puskesmas Purwakarta reach.

After data selection process, data processing and transformation finished, it will produce a table, which is ready to classify as decision tree form using C4.5 algorithm, be described below.

Table 2.2. Table of case

Nama Penyakit	Umur	Wilayah
Diare	45 - 55 tahun	Sindangkasih
Diare	29 hari - 1 tahun	Nagri Tengah
Tuberkolosis	5 - 9 tahun	Luar Wilayah
Tuberkolosis	5 - 9 tahun	Luar Wilayah
Pneumonia	29 hari - 1 tahun	Sindangkasih
Diare	1 - 4 tahun	Luar Wilayah
Diare	5 - 9 tahun	Luar Wilayah
ISPA	1 - 4 tahun	Sindangkasih
ISPA	1 - 4 tahun	Sindangkasih
ISPA	1 - 4 tahun	Luar Wilayah
ISPA	1 - 4 tahun	Luar Wilayah
Dispepsia	20 - 44 tahun	Luar Wilayah
Dispepsia	45 - 55 tahun	Luar Wilayah
ISPA	5 - 9 tahun	Nagri Kidul
ISPA	5 - 9 tahun	Sindangkasih
ISPA	10 - 14 tahun	Sindangkasih
Gangguan Saraf	10 - 14 tahun	Nagri Kidul
ISPA	10 - 14 tahun	Nagri Kidul
Tukak Lambung	20 - 44 tahun	Luar Wilayah
Demam Tifoid	1 - 4 tahun	Sindangkasih
Demam Tifoid	5 - 9 tahun	Sindangkasih
ISPA	15 - 19 tahun	Luar Wilayah
Diabetes Melitus	56 - 59 tahun	Nagri Kidul
Diabetes Melitus	60 - 69 tahun	Luar Wilayah
Hipertensi Primer	20 - 44 tahun	Luar Wilayah
Hipertensi Primer	20 - 44 tahun	Luar Wilayah
Hipertensi Primer	45 - 55 tahun	Sindangkasih
Gangguan Kulit	20 - 44 tahun	Luar Wilayah
Cacar Air	20 - 44 tahun	Nagri Kidul
Common Cold	29 hari - 1 tahun	Sindangkasih

2.7.4 Data Mining

Data mining is a explore-information process such us a disease classification based on region with form of decision tree.

Before constructing the decision tree, the data above is processed using C4.5 algorithm by calculating the entropy and choosing an attribute as the root based on the highest gain.

This is calculating process for the entropy.

Entropy total = $((-\frac{5}{30} * \log_2(\frac{5}{30})) + ((-\frac{1}{30} * \log_2(\frac{1}{30}))$ + $((-\frac{10}{30} * \log_2(\frac{10}{30})) + ((-\frac{14}{30} * \log_2(\frac{14}{30})) = 1{,}6358$

1. Atribut Nama Penyakit

- *Entropy* Cacar Air = $((-\frac{1}{1} * \log_2(\frac{1}{1})) = 0$
- *Entropy* Common Cold = $((-\frac{1}{1} * \log_2(\frac{1}{1})) = 0$
- *Entropy* Demam Tifoid = $((-\frac{2}{2} * \log_2(\frac{2}{2})) = 0$
- *Entropy* Diabetes Melitus = $((-\frac{1}{2} * \log_2(\frac{1}{2})) + ((-\frac{1}{2} * \log_2(\frac{1}{2})) = 1$
- *Entropy* Diare = $((-\frac{1}{4} * \log_2(\frac{1}{4})) + ((-\frac{1}{4} * \log_2(\frac{1}{4})) + ((-\frac{2}{4} * \log_2(\frac{2}{4})) = 1{,}5$
- *Entropy* Dispepsia = $((-\frac{2}{2} * \log_2(\frac{2}{2})) = 0$
- *Entropy* Gangguan Kulit = $((-\frac{1}{1} * \log_2(\frac{1}{1})) = 0$
- *Entropy* Gangguan Saraf = $((-\frac{1}{1} * \log_2(\frac{1}{1})) = 0$
- *Entropy* Hipertensi Primer = $((-\frac{1}{3} * \log_2(\frac{1}{3})) + ((-\frac{2}{3} * \log_2(\frac{2}{3})) = 0{,}9183$
- *Entropy* ISPA = $((-\frac{2}{9} * \log_2(\frac{2}{9})) + ((-\frac{4}{9} * \log_2(\frac{4}{9})) + ((-\frac{3}{9} * \log_2(\frac{3}{9})) = 1{,}5305$
- *Entropy* Pneumonia = $((-\frac{1}{1} * \log_2(\frac{1}{1})) = 0$
- *Entropy* Tuberkolosis = $((-\frac{2}{2} * \log_2(\frac{2}{2})) = 0$
- *Entropy* Tukak Lambung = $((-\frac{1}{1} * \log_2(\frac{1}{1})) = 0$

2. Atribut Umur

- *Entropy* 29 hari – 1 tahun = $((-\frac{1}{3} * \log_2(\frac{1}{3})) + ((-\frac{2}{3} * \log_2(\frac{2}{3})) = 0{,}9183$
- *Entropy* 1 – 4 tahun = $((-\frac{3}{6} * \log_2(\frac{3}{6})) + ((-\frac{3}{6} * \log_2(\frac{3}{6})) = 1$
- *Entropy* 5 – 9 tahun = $((-\frac{1}{6} * \log_2(\frac{1}{6})) + ((-\frac{2}{6} * \log_2(\frac{2}{6})) + ((-\frac{3}{6} * \log_2(\frac{3}{6})) = 1{,}4591$
- *Entropy* 10 – 14 tahun = $((-\frac{2}{3} * \log_2(\frac{2}{3})) + ((-\frac{1}{3} * \log_2(\frac{1}{3})) = 0{,}9183$
- *Entropy* 15 – 19 tahun = $((-\frac{1}{1} * \log_2(\frac{1}{1})) = 0$
- *Entropy* 20 – 44 tahun = $((-\frac{1}{6} * \log_2(\frac{1}{6})) + ((-\frac{5}{6} * \log_2(\frac{5}{6})) = 0{,}6500$
- *Entropy* 45 – 55 tahun = $((-\frac{2}{3} * \log_2(\frac{2}{3})) + ((-\frac{1}{3} * \log_2(\frac{1}{3})) = 0{,}9183$
- *Entropy* 56 – 59 tahun = $((-\frac{1}{1} * \log_2(\frac{1}{1})) = 0$
- *Entropy* 60 – 69 tahun = $((-\frac{1}{1} * \log_2(\frac{1}{1})) = 0$

After calculating the entropy, next step is calculating the gain for name of disease attribute and age attribute.

1. *Gain* Nama Penyakit

$1{,}6358 – ((\frac{1}{30} * 0) + (\frac{1}{30} * 0) + (\frac{2}{30} * 0) + (\frac{2}{30} * 1) + (\frac{4}{30} * 1{,}5) + (\frac{2}{30} * 0) + (\frac{1}{30} * 0) + (\frac{1}{30} * 0) + (\frac{3}{30} * 0{,}9183) + (\frac{19}{30} * 1{,}5305) + (\frac{1}{30} * 0) + (\frac{2}{30} * 0) + (\frac{1}{30} * 0) = 0{,}8182$

2. *Gain* Umur

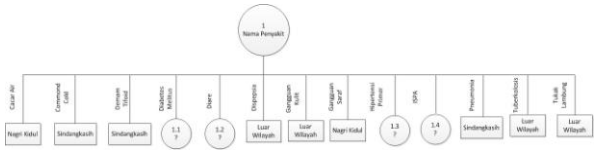
$1{,}6358 – ((\frac{3}{30} * 0{,}9183) + (\frac{6}{30} * 1) + (\frac{6}{30} * 1{,}4591) + (\frac{3}{30} * 0{,}9183) + (\frac{1}{30} * 0) + (\frac{6}{30} * 0{,}6500) + (\frac{3}{30} * 0{,}9183) + (\frac{1}{30} * 0) + (\frac{1}{30} * 0) = 0{,}7385$

The picture below is a result of entropy and gain calculation in table form.

Nama Atribut	Nilai Atribut	Jml Kasus (total)	Nagri Kidul	Nagri Tengah	Sindangkasih	Luar Wilayah	Entropy	Gain
Total		30	5	1	10	14	1,6358	
Nama Penyakit								0,8182
	Cacar Air	1	1	0	0	0	0	
	Common Cold	1	0	0	1	0	0	
	Demam Tifoid	2	0	0	2	0	0	
	Diabetes Melitus	2	1	0	0	1	1	
	Diare	4	0	1	1	2	1,5	
	Dispepsia	2	0	0	0	2	0	
	Gangguan Kulit	1	0	0	0	1	0	
	Gangguan Saraf	1	1	0	0	0	0	
	Hipertensi Primer	3	0	0	1	2	0,9183	
	ISPA	9	2	0	4	3	1,5305	
	Pneumonia	1	0	0	1	0	0	
	Tuberkolosis	2	0	0	0	2	0	
	Tukak Lambung	1	0	0	0	1	0	
Umur								0,7385
	29 hari - 1 tahun	3	0	1	2	0	0,9183	
	1 - 4 tahun	6	0	0	3	3	1	
	5 - 9 tahun	6	1	0	2	3	1,4591	
	10 - 14 tahun	3	2	0	1	0	0,9183	
	15 - 19 tahun	1	0	0	0	1	0	
	20 - 44 tahun	6	1	0	0	5	0,6500	
	45 - 55 tahun	3	0	0	2	1	0,9183	
	56 - 59 tahun	1	1	0	0	0	0	
	60 - 69 tahun	1	0	0	0	1	0	

Picture 2.1. Table picture of entropy and gain calculation

The table above shows that the highest gain is Name of Disease with amount 0.8182. Then, the Name of Disease (Nama Penyakit) can be the first node and be the root node. There are 13 attributes in Name of Disease. Those attributes still need more calculation for Diabetes Mellitus, Diare, Hipertensi Primer and ISPA because they have not been in any class. The result from the table above can be illustrated on decision tree such as this picture.



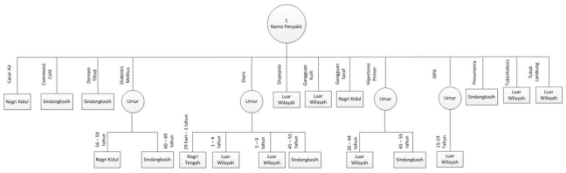
Picture. 2.2. First node of decision tree form

The final entropy and gain calculation and decision tree can be seen on the table picture below.

Nama Atribut	Nilai Atribut	Jml Kasus (total)	Nagri Kidul	Nagri Tengah	Sindangkasih	Luar Wilayah	Entropy	Gain
Nama Penyakit - ISPA		9	2	0	4	3	1,5305	
Umur								0,6416
	29 hari - 1 tahun	0	0	0	0	0	0	
	1 - 4 tahun	4	0	0	2	2	1	
	5 - 9 tahun	2	1	0	1	0	1	
	10 - 14 tahun	2	1	0	1	0	1	
	15 - 19 tahun	1	0	0	0	1	0	
	20 - 44 tahun	0	0	0	0	0	0	
	45 - 55 tahun	0	0	0	0	0	0	
	56 - 59 tahun	0	0	0	0	0	0	
	60 - 69 tahun	0	0	0	0	0	0	

Picture 2.3. Table picture of entropy and gain for final calculation

On the table picture, there are attributes that have not entered any classes. But, because the only attribute used is age attribute, there will be no more calculation for them. Thus, the calculation ends here.



Picture 2.4. Final form of decision tree

2.7.5 Interpretation / Evaluation

Based on the decision tree above, the all cases have entered to the classes. Therefore, the decision tree above can be the material for Puskesmas Purwakarta to decide the extension to certain region. The decision tree will be served on a table below.

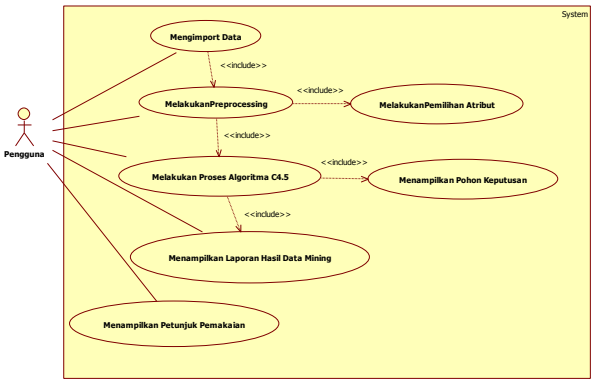
Table. 2.3. Table of decision tree result

Nama Penyakit	Keterangan dan Wilayah
Cacar Air	Berada di Nagri Kidul
Commond Cold	Berada di Sindangkasih
Demam Tifoid	Berada di Sindangkasih
Diabetes Melitus	Umur 56 – 59 tahun berada di Nagri Kidul Umur 60 – 69 tahun berada di Sindangkasih
Diare	Umur 29 hari – 1 tahun berada di Nagri Tengah Umur 1 – 4 tahun bearda di Luar Wilayah Umur 5 – 9 tahun

	bearda di Luar Wilayah Umur 55 – 55 tahun bearda di Sindangkasih
Dispepsia	Berada di Luar Wilayah
Ganggua Kulit	Berada di Luar Wilayah
Gangguan Saraf	Berada di Nagri Kidul
Hipertensi Primer	Umur 20 – 44 tahun bearda di Luar Wilayah Umur 45 – 55 tahun bearda di Sindangkasih
ISPA	Umur 15 – 19 tahun berada di Luar Wilayah
Pneumonia	Berada di Sindangkasih
Tuberkulosis	Berada di Luar Wilayah
Tukak Lambung	Berada di Luar Wilayah

2.8 Functional Requirement Analysis

2.8.1 Use Case Diagram



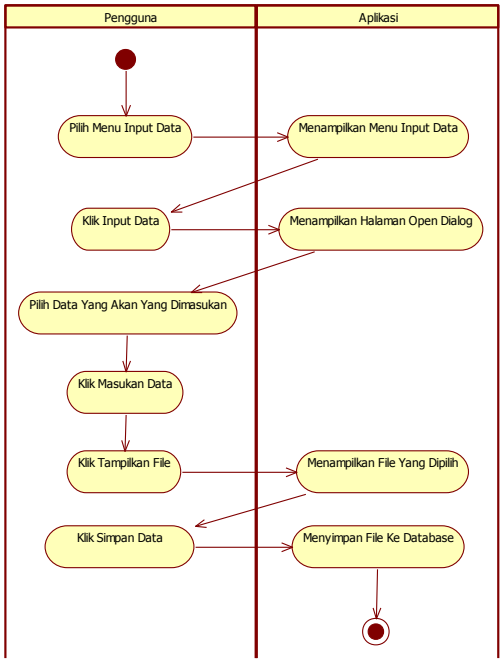
Picture 2.5. Use case diagram data mining application

2.8.2 Skenario use case input data

Use Case Name	**Input Data**	
Goal In Context	Masuk ke dalam aplikasi dan data bisa dimasukan ke dalam aplikasi dan *database*	
Precondition	-	
Succesful End Condition	Mengakses menu input data, menampilkan data ke dalam aplikasi dan berhasil memasukan data ke *database*	
Fail End Condition	Tidak bisa mengakses menu input data, tidak bisa menampilkan data ke dalam aplikasi dan tidak bisa memasukan data ke *database*	
Primary Actor	*User*	
Secondary Actor	-	
Trigger	Mengklik tombol simpan data	
Main Flow	Step	Action
	1.	Masuk ke dalam aplikasi
	2.	Memilih menu Input Data
	3.	Aplikasi akan menampilkan menu input data
	4.	*User* akan memilih data yang berformat Microsoft Excel
	5.	Setelah data terpilih, *user* akan memasukan data ke dalam aplikasi dan akan ditampilkan ke dalam aplikasi
	6.	Data yang sudah terpilih tersebut akan disimpan ke dalam *database* dengan mengklik tombol simpan data
Extension	-	-

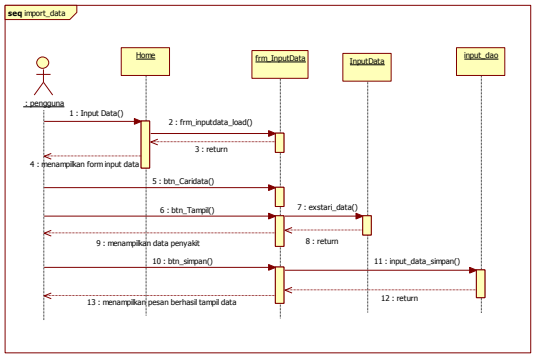
Picture 2.6. Table picture of use case input data

2.8.3 Activity Diagram Input Data



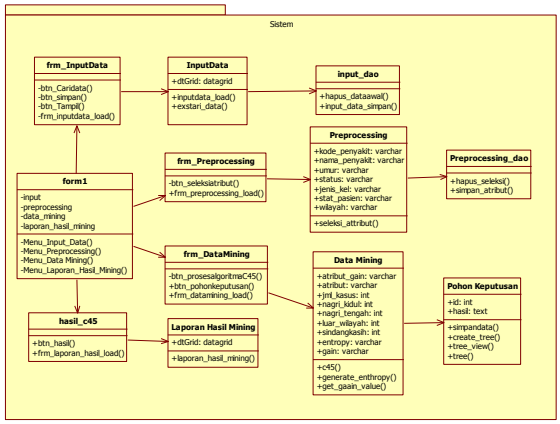
Picture 2.7. Activity diagram input data

2.8.4 Sequence Diagram Input Data



Picture 2.8. Sequence Diagram Input Data

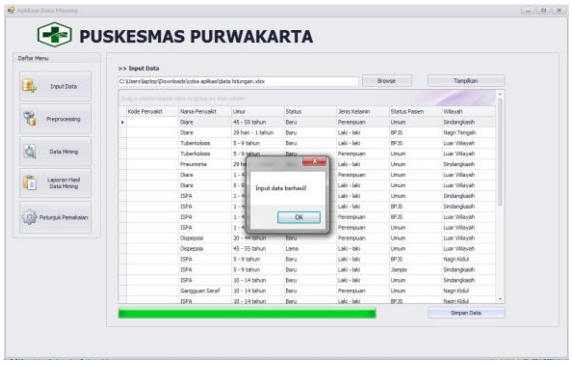
2.8.5 Class Diagram



Picture 2.9. Class Diagram

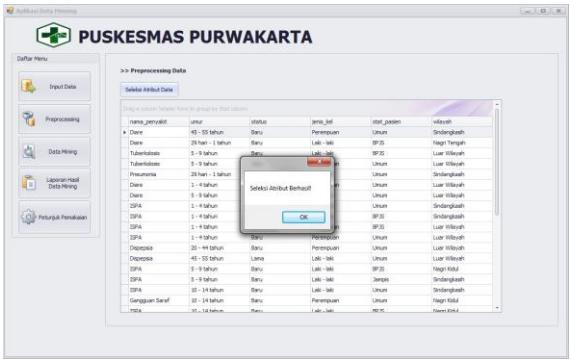
2.9 Interface Implementation

a. Display Input Data Menu



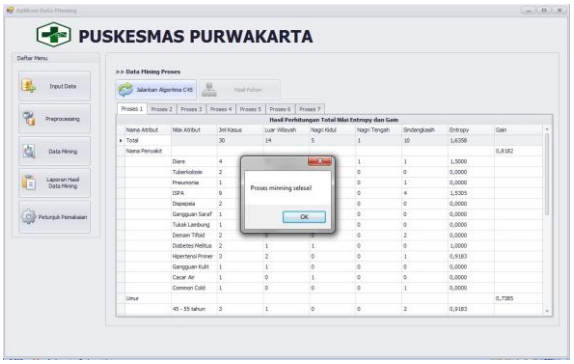
Picture 2.10. Display input data menu

b. Display Preprocessing Menu



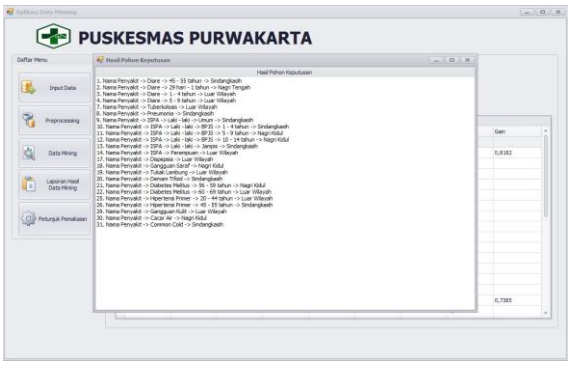
Picture 2.11. Display preprocessing menu

c. Display Data Mining Menu



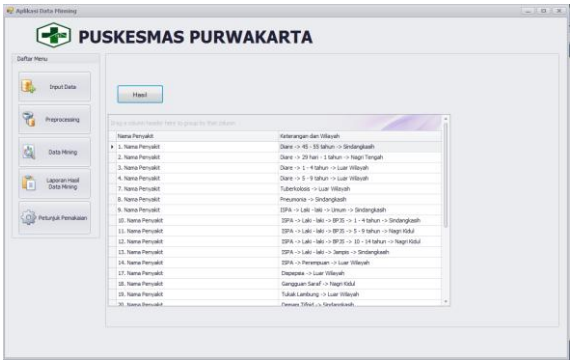
Picture 2.12. Display data mining menu

d. Display Decision Tree



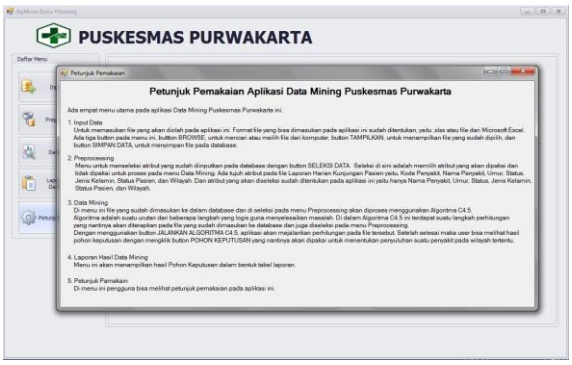
Picture 2.13. Display decision tree

e. Data Mining Result Report



Picture 2.14. Data Mining Result Report

f. How To Use



Picture 2.15. How tu use

3. CONCLUSION

Based on the analysis result and the trial on application Data Mining system with Classification method and C4.5 Algorithm, the conclusion is :

The application can help Puskesmas Purwakarta for their extension about disease to the community be more effective according to daily visit reports processed by the application. The information produced by the application is a decision tree. It can be a reference to decide which region that will be held an extension.

THANKS TO

The writer realizes that many problems on making this research. But, the help, guidance, and cooperation given from many side and bless given from Allah SWT, the writer can solve all the problems. Thus, the writer really wants to say a bunch thanks to :

1. Allah SWT for his blesses and favor.
2. Prophet Muhammad SAW for his amorousness.
3. To my parents, Siti Nurjanah and the late Suratno, who always give me courage, advice and countless attention.
4. To Mr. Alif Finandhita, S.Kom., M.T, as a thesis adviser who always gives me courage and guidance during composing this research.

5. To Mrs. Dian Dharmayanti, S.T., M.Kom. as a thesis reviewer who has given me advice and critic during composing this researcH.
6. To Mrs. Gentisya Tri Mardiani, S.Kom., M.Kom. as my academic advisor and thesis examiner who has given me advice and critic during composing this research.
7. Mrs. Anne Hediana Koesmah, HJ., dr., MM as Head of Health Department at Purwakarta Regency who gave me a permission for doing a research in her field.
8. Mrs Hj. Atik Hayati, AmKeb, MMKes as Head of Puskesmas Purwakarta who gave me permission to do my research at Puskesmas Purwakarta.
9. To all my campus mate at Unikom especially IF-14 class 2011 who always support and share their opinions during composing the research.
10. And thanks to all people that cannot be mentioned here.

REFERENCES

[1] Larose, Daniel T. 2005. *Discovering Knowledge in Data : An Introduction to Data Mining.* John Willey & Sons, Inc.
[2] Kusrini dan Luthfi, Emha Taufiq. 2009. *Algoritma Data Mining*. Penerbit Andi. Yogyakarta.
[3] Adi Swara, Jaka. *Aplikasi Data Mining Menggunakan Metode Decison Tree Untuk Pemantauan Distribusi Penjualan Sepeda Motor Di PD. Wijaya Abadi Bandung.*
[4] Sommerville, Ian. 2011. *Software Engineering : 9th Edition.* Addison-Wesley. Boston Massachusetts.
[5] Sumanthi, S. Sivanandam, S.N. 2009. *Introduction to Data Mining and its Applications*. Spinger. Verlag Berlin Heidelberg.
[6] Kadir, Abdul. 2003. *Pengenalan Sistem Informasi.* Penerbit Andi Yogyakarta. Yogyakarta.
[7] Sutabri, Tata. 2005. *Sistem Informasi Manajemen.* Penerbit Andi Yogyakarta. Yogyakarta.
[8] Fathansyah. 2004. *Buku Teks Komputer Basis Data*. Informatika Bandung.
[9] Hariyanto, Bambang. 2004. *Sistem Manajemen Basis Data : Pemodelan, Perancangan, dan Terapannya.* Informatika Bandung.
[10] Connolly, T. Begg, C. 2010. *Database Systems : A Practical Approach To Design, Implementation, and Management. 5th Edition*. America : Pearson Education.
[11] Utdirartatmo, Firar. 2002. *Mengelola Database Server MySQL Di Linux Dan Windows*. Penerbit Andi Yogyakarta.
[12] Weisfeld, Matt. 2009. *Object-Oriented Thought Process, Third Edition*. Addison-Wesley
[13] Fowler, Martin. 2005. *UML Distilled. Panduan Singkat Bahasa Permodelan Objek Standar Edisi 3*. Penerbit Andi Yogyakarta. Yogyakarta.
[14] Presman, Roger S. 1997. *Software Engineering : A Practioner's Approach 4th Edition*. Mcgraw Hill.