Implementation of the C4.5 Algorithm in Predicting the Number of Outpatient Visits Using JKN-KIS at Noongan Hospital

Liza Wikarsa*[1], Vivie Deyby Kumenap[2], Kevin Kristi Toar[3]
[1,2,3]Universitas Katolok De La Salle Manado; Kairagi I Kombos Manado, Belakang Wenang Permai II, Manado, 95000, Telp: 0431-813148, Fax: 0431-813161
[1,2,3]Program Studi Teknik Informatika, Fakultas Teknik, UDLSM, Manado
e-mail: ***1wikarsa@unikadelasalle.ac.id**, [2]vkumenap@unikadelasalle.ac.id, [3]14013053@unikadelasalle.ac.id

Abstract

Since 2013, the government has issued the National Health Insurance (JKN) program through the Social Security and Health Administration (BPJS Kesehatan) to provide social and health insurance services. JKN participants will get a Healthy Indonesia Card (KIS) to ease the burden of medical expenses at the hospital. During the pandemic of Covid-19, Noongan Hospital was included as one of the referral hospitals for COVID-19 patients for nearby hospitals and health centers with a coverage of the Southeast Minahasa district, North Sulawesi. Noticeably, 70% of its patients use the JKN-KIS card to get health treatments and more than half of the patients are outpatients. To anticipate the number of outpatients visits using JKN-KIS, a web-based application was built to generate a predictive model using the C4.5 algorithm. The performance of this predictive model has a classification accuracy of 91,7% and both precision and recall of 95%. The number of outpatient visits using JKN-NIS has increased by 83,33% since the pandemic of Covid-19. Examination flow, medical check-ups, queue length, doctor's expertise, and health treatment objectives influence outpatient visits. This predictive model provides future insights for the hospital management to rationally allocate healthcare resources and improve the efficiency of outpatient services.

Keywords—3 Health Treatments, C4.5 Algorithm, Prediction, Covid-19

1. INTRODUCTION

The Social Security and Health Administration (BPJS Kesehatan) is an autonomous public legal entity that was established to support the government's program, namely the National Health Insurance (JKN) by issuing the Healthy Indonesia Card (KIS). The government through BPJS Kesehatan aids every citizen and foreigner who has lived at least six months in Indonesia to have social and health insurance. In this JKN program, there are several classes, namely first class, second class, and third class which have differences in the monthly fee and the facilities covered. Class 1 is the highest in this program with the best facilities that the hospital can provide to first-class JKN participants starting from first-class type treatment rooms to services. Then followed by Class 2 and Class 3 whose monthly contributions are less. The government itself provides subsidies for the underprivileged so that they can enjoy the third-class JKN. Whilst, civil servants or civil servants are given allowances in the form of second-class JKN according to their class or rank [1,2]. In addition, KIS users can also get class facilities above them by paying the difference from those registered in JKN with the fees to be paid.

During this pandemic period, several hospitals were used as referral hospitals for COVID-19 patients in North Sulawesi where Noongan Hospital was included in the list of referral hospitals. The Noongan Regional General Hospital (Noongan Hospital) is a hospital located in the village of Noongan I, West Langowan District, Minahasa Regency, North Sulawesi Province. This hospital is a referral hospital for nearby hospitals and health centers covering the Southeast Minahasa district. More than 70 percent of patients who come to check themselves at this hospital use the JKN-KIS card to get a fee waiver during the examination process [3]. In addition, more than half the number of patients who come to this hospital are outpatients who come to check themselves with some complaints or do a medical check-up. Currently, the BPJS section only operates every Monday to Friday. The patient's need for affordable examination fees is the main factor that makes patients come to check themselves using the JKN-KIS card. Also, several factors influence patients to go for a check-up at this hospital, such as the distance between their residence and the Noongan Hospital, the services of the BPJS hospital staff, available medical facilities, and the availability of doctors on duty there.

To anticipate the number of outpatients visits, Noongan Hospital needs to build a prediction model that can provide pertinent data for the hospital's management not only to allocate healthcare resources rationally but also to increase the number of outpatient visits by improving the efficiency of outpatient services. In this study, C4.5 is used to build the prediction model for Noongan Hospital as this algorithm is commonly used in data mining as a Decision Tree Classifier to generate a decision based on a certain sample of data that is outpatient visits using JKN-KIS for this research. C4.5 itself is a development of the previous similar algorithm, namely the ID3 algorithm with several improvements such as resolved missing values, continuous data processing, and pruning. This algorithm utilizes several variables (multivariate predictors) that can affect an event to get the desired result [4,5].

C4.5 algorithm is proven useful for prediction in certain cases [4,6,7]. Haruna et al. used the C4.5 algorithm to diagnose coronary artery disease [4]. The parameters used for the diagnosis included blood pressure, glucose, cholesterol, triglyceride, high density of lipoprotein (HDL), low density of lipoprotein (LDL), creatinine, body mass index (BMI), heart rate, chest pain, and diagnostic CAD. This research dataset was taken from Murtala Muhammad General Hospital and Abdullahi Wase General Hospitals in Kano State in Nigeria. The research used 506 datasets that were collected from 2003 to 2017. The accuracy of diagnosing this coronary artery disease using the C4.5 algorithm was 94,27%. Another research was conducted by Andriani concerning the classification based on the C4.5 algorithm for detection of an increase in the case fatality rate of Dengue fever [6]. It used the CFR Dengue dataset in all provinces in Indonesia to find the most influencing factors to cause an increase in this disease in a province. Based on evaluation results with Confusion Matrix and ROC curve, applying classification rules to training data and data testing obtained an average accuracy value of 80.245%, while the accuracy value shown by the ROC curve is 0.913. Meanwhile, Noviandi used the C4.5 algorithm to predict the diagnosis of diabetes mellitus in patients who have just given birth. The independent variables used in this research were the number of patients, sugar levels, blood pressure, insulin, body mass index, and age. The results showed this C4.5 algorithm could accurately predict as much as 70.32% with 9 rules by which 4 rules with no/negative class, and 5 rules with yes/positive class [7].

This research would build an application that can predict the number of outpatient visits using JKN-KIS at Noongan Hospital. The algorithm for prediction used was the C4.5 algorithm with eight independent variables such as medical check-ups, referral, an extension of BPJS service, queue length, health treatment objectivities, doctor's expertise, examination flow, and hospital service satisfaction. The contribution of this research is to build a prediction model to help provide data for the management and decision-makers at Noongan Hospital about the future trend of the number of outpatients visits using JKN-KIS. This will certainly have impacts on the allocation of healthcare resources, the efficiency of outpatient service, and more. Lastly, the organization of this paper is divided into five parts, namely introduction, research method, results and discussion, conclusion, and future works.

2. RESEARCH METHOD

2.1 Research Participants

Based on the recap of JKN-KIS patient visits at the Noongan Hospital carried out every three months, this research used the data from March 2021 that had a total of three hundred and two (302) JKN-KIS outpatient visits. Using table Isaac and Michael [8], the research sample consisted of 161 participants for this research by which they were asked to fill in the online questionnaires using Google Form. The significance level for this research is 0.05.

2.2 Research Framework

The research framework consisted of four stages, namely data collection, pre-processing, processing, and validation. The following sub-section will briefly explain what is to be done and the expected outputs for each stage.

2.2.1 Stage I: Data Collection

In this stage, research questions were posted in Google Form for the participants to fill in. The purpose of this questionnaire is to identify and analyze the responses of the participants with regards to medical checkups, referral, an extension of BPJS service, queue length, health treatment objectivities, doctor's expertise, examination flow, and hospital service satisfaction at Noongan Hospital.

2.2.2 Stage II: Pre-processing

After collecting the data from the previous stage, these data must be further processed using several statistical tests, namely the validity and correlation of the data using the Pearson formula, reliability testing using the Cronbach's Alpha formula, and the T-test. These tests were performed on the SPSS 25 (Statistical Product and Service Solutions) program to ensure that questionnaire data is feasible to use. Data that has passed the statistical test will be used at the processing stage.

2.2.3 Stage III: Processing

In this stage, a web-based application was built to predict the number of outpatient visits using JKN-KIS at Noongan Hospital. This application implemented the C4.5 algorithm using PHP and MySQL for the database. The steps of the application of C4.5 are as follows [3]:

1. *Preparation of training data*: the dataset contained the responses of the 161 participants pertinent to the eight independent variables.
2. *Calculation of the tree root*: calculation is done by calculating the **entropy value** followed by calculation of **the gain value of each attribute**. Attribute with value **the**

highest gain ratio will be the root of the tree. Formula entropy calculation can be seen in equation (1).

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

Note:
S = set of cases
n = Number of partitions S
Pi = Proportion of Si to S

The above equation shows the formula for calculating the entropy of a set S where S is the value of the attributes used in the dataset. While n is the number of partitions of the set S in this case there are two partitions, namely "positive" and "negative". The calculation of **the information gain (gain) value** is done with the following formula:

$$Gain\ (S, A) = entropy\ (S) - \Sigma_{i=1}^{n} \frac{|s_i|}{s} * Entropy\ (Si) \quad (2)$$

Note:
S = Set of Cases
A = Feature/Attribute
n = Number of attribute partitions A
|Si| = Proportion of Si to S
|S| = Number of cases in S

Next to get the value of the **Split Info** of the attribute A by which equation for the split info is as follows:

$$SplitInfo\ (S, A) = - \Sigma_{i=1}^{n} \frac{s_i}{s} \log_2 \frac{s_i}{s} \quad (3)$$

Lastly, it is to find the **Gain Ratio** that is "… used as a classification attribute to reclusively build the branches of the tree from dataset until the complete structure of the tree is formed" [2, pp.49]. The equation for the gain ratio (4) is provided below.

$$GainRatio(S, A) = \frac{Gain(S,A)}{SplitInfo(S,A)} \quad (4)$$

Note:
Gain (S,A) = Information Gain on attribute A
SplitInfo(S,A) = Split Info on attribute A

The results of the entropy calculation are used for the calculation of gain based on equation (2). After that, one needs to calculate the split info using equation (3). Lastly, the gain ratio is calculated by dividing the value of information gain on attribute A by the value of split info on attribute A as shown in equation (4).

Attribute with the highest gain ratio will become the root of the decision tree or root, while other attributes that are not attributed with the highest gain ratio will be recalculated. From the result, the recalculation will obtain the highest gain ratio back which is then used as a sub root under the root. Then the calculation process as before is repeated back until all selected attributes become sub root so that all the attributes in the partitioned dataset form a decision tree.

3. *Repeat the calculations of entropy, information gain, split info and gain ratio on attributes unpartitioned.*

4. *The calculation process above is repeated until all partitioned attributes*. This process will stop when all records in node N get the same class same, with no unpartitioned attributes, and nothing recorded in an empty branch.

2.2.4 Stage IV: Validation

Confusion Matrix is used to display accuracy values in the form of percentage figures for the accuracy of the classification rule in classifying data equipped with precision values or confidence and recall values or sensitivity [9].

3. RESULTS AND DISCUSSION

This section would address the results of this research in accord with the research framework mentioned in the previous section.

3.1 Data Collection

The research data was collected using Google Form and filled in by 161 participants who live in North Sulawesi. These data contained the responses of the participants about a medical check-up, referral, extension of BPJS service, queue length, health treatment objectivities, doctor's expertise, examination flow, and hospital service satisfaction. Table 1 enlisted all the variables, indicators, and scoring used in this research.

Table 1 Variables, Indicators, and Scoring

Variables	Indicators	Score	Total Responses
Medical check-ups	Very unwilling	4	26
	Not willing	3	50
	Willing	2	79
	Very willing	1	5
Referral	Very difficult	4	31
	Difficult	3	75
	Easy	2	52
	Very easy	1	3
Extension of BPJS service	Very difficult	4	36
	Difficult	3	73
	Easy	2	49
	Very easy	1	3
Queue length	Very long	4	46
	Long	3	51
	Short	2	20
	Very short	1	44
Health treatment objectives	Very objective	4	80
	Objective	3	30
	Not Objective	2	29
	Very not objective	1	22
Doctor's expertise	Excellent	4	35
	Good	3	69
	Moderate	2	38
	Not good	1	19
Examination flow	Very difficult	4	46
	Difficult	3	53
	Easy	2	23
	Very easy	1	39

Variables	Indicators	Score	Total Responses
Hospital service satisfaction	Very unsatisfied	4	41
	Not satisfied	3	72
	Satisfied	2	42
	Very satisfied	1	6

3.2 Pre-processing

The reliability and validity of the questionnaires for the C4.5 model in this research were tested using SPSS 25 (Statistical Product and Service Solutions) program. Based on the reliability test on the eight items in the questionnaires, the Cronbach's alpha value is 0.891 which exceeds the alpha value of 0.6, hence the questionnaire can be declared reliable [11,12]. The validity of the questionnaires was between 0,493 and 0,877 (r-count) and the r-table was 0,1543 as shown in 2. The value of the r-Count is greater than the r-table and therefore the items in the questionnaire are declared valid. In addition, the value of Sig. (2-tailed) is less than 0.05 and the Pearson Correlation is positive, then the questionnaire item is valid [12].

Table 2 Validity Test

Variable	r-Count	r-Table	*Sig.(2-tailed)*	Validity
Medical check up	0,551	0,1543	0	Valid
Referral	0,513	0,1543	0	Valid
Extension of BPJS service	0,833	0,1543	0	Valid
Queue length	0,877	0,1543	0	Valid
Health treatment objectives	0,493	0,1543	0	Valid
Doctor's expertise	0,811	0,1543	0	Valid
Examination flow	0,785	0,1543	0	Valid
Hospital service satisfaction	0,842	0,1543	0	Valid

3.3 Processing

Processing on the dataset was performed using the C4.5 algorithm as follows.

1. *Preparation of training data*: the dataset used came from the responses of the 161 participants pertinent to the eight independent variables. The dataset is split into 70% training data (113 data) and 30% testing data (48 data).
   The calculation begins by setting the conditions for data to be labeled as Up (increasing) or Down (decreasing) as shown in Table 3. In this case, a condition is made that the average value of each questionnaire with a score below 2.6 will be labeled down while scores of 2.6 and greater than 4 will be labeled up.

Table 3 Data Labelling

Up	Down	Total
63	50	113

2. *Calculation of the tree root.*
   Calculate the entropy value followed by calculating the gain value of each attribute as demonstrated in Table 4. Then, one needs to calculate the split info dan lastly gain ratio. Attribute with value the highest gain ratio will be the root of the tree.

Table 4 Entropy Value of Total Attributes

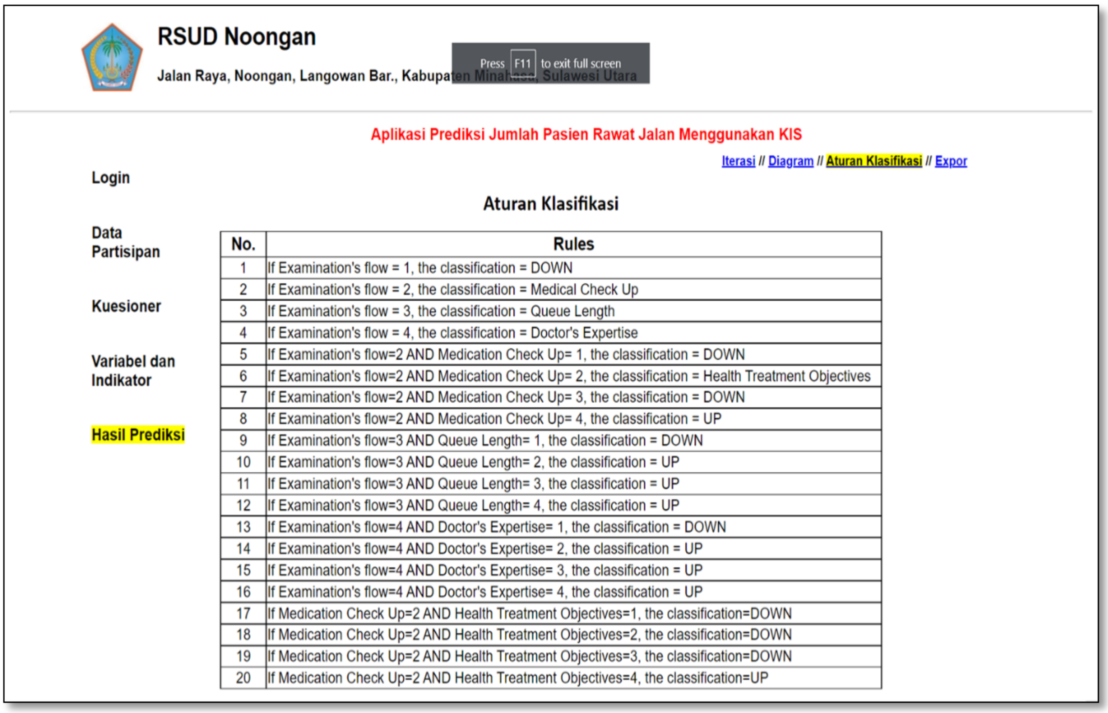
Total	Up	Down	Entropy
113	63	50	0.990432

Table 5 The Entropy Information Gain, Split Info, and Gain Ratio of Each Attribute

Tree Root		Total	Up	Down	Entropy	Information Gain	Split Info	Gain Ratio
Total		*113*	*68*	*45*	*0.990432*			
Medical Check-ups	1	6	0	6	0	0.149092	1.696024	0.087907
	2	56	29	27	0.99908			
	3	26	12	14	0.995727			
	4	25	22	3	0.529361			
Referral	1	3	0	3	0	0.061851	1.677154	0.036878
	2	36	17	19	0.997772			
	3	48	33	14	0.89011			
	4	26	13	13	1			
Extension of BPJS Service	1	3	1	2	0.918296	0.068952	1.697285	0.040625
	2	36	13	23	0.943602			
	3	45	32	13	0.867282			
	4	29	17	12	0.978449			
Queue Length	1	37	0	37	0	0.714744	1.961338	0.364417
	2	19	8	11	0.981941			
	3	27	26	1	0.228538			
	4	30	29	1	0.210842			
Health Treatment Objectives	1	19	0	19	0	0.627733	1.784758	0.351719
	2	26	2	24	0.391244			
	3	13	10	3	0.77935			
	4	55	51	4	0.37602			
Doctor's Expertise	1	19	0	19	0	0.641302	1.90125	0.337306
	2	28	3	25	0.491237			
	3	46	40	6	0.558629			
	4	20	20	0	0			
Examination Flow	1	33	0	33	0	0.75324	1.970116	0.382333
	2	19	4	15	0.742488			
	3	32	31	1	0.200622			
	4	29	28	1	0.216397			
Hospital Service Satisfaction	1	6	0	6	0	0.077328	1.7268	0.044781
	2	24	12	12	1			
	3	52	34	18	0.930586			
	4	31	17	14	0.993234			
							The Highest Gain Ratio	0.382333

Table 5 showed the calculation of entropy, information gain, split info, and gain ratio. It was found that the root of the decision tree model for this research is the examination flow due to its highest gain ratio of 0.382333.

3. *Repeat entropy and gain calculations on attributes unpartitioned.*
4. *The calculation process above is repeated until all partitioned attributes*



No.	Rules
1	If Examination's flow = 1, the classification = DOWN
2	If Examination's flow = 2, the classification = Medical Check Up
3	If Examination's flow = 3, the classification = Queue Length
4	If Examination's flow = 4, the classification = Doctor's Expertise
5	If Examination's flow=2 AND Medication Check Up= 1, the classification = DOWN
6	If Examination's flow=2 AND Medication Check Up= 2, the classification = Health Treatment Objectives
7	If Examination's flow=2 AND Medication Check Up= 3, the classification = DOWN
8	If Examination's flow=2 AND Medication Check Up= 4, the classification = UP
9	If Examination's flow=3 AND Queue Length= 1, the classification = DOWN
10	If Examination's flow=3 AND Queue Length= 2, the classification = UP
11	If Examination's flow=3 AND Queue Length= 3, the classification = UP
12	If Examination's flow=3 AND Queue Length= 4, the classification = UP
13	If Examination's flow=4 AND Doctor's Expertise= 1, the classification = DOWN
14	If Examination's flow=4 AND Doctor's Expertise= 2, the classification = UP
15	If Examination's flow=4 AND Doctor's Expertise= 3, the classification = UP
16	If Examination's flow=4 AND Doctor's Expertise= 4, the classification = UP
17	If Medication Check Up=2 AND Health Treatment Objectives=1, the classification=DOWN
18	If Medication Check Up=2 AND Health Treatment Objectives=2, the classification=DOWN
19	If Medication Check Up=2 AND Health Treatment Objectives=3, the classification=DOWN
20	If Medication Check Up=2 AND Health Treatment Objectives=4, the classification=UP

Figure 1 The classification rules

Figure 1 enlisted the twenty rules that determined whether the number of outpatient visits using JKN-KIS at Noongan Hospital experienced an increase or a decrease during the pandemic of Covid-19. Also, these rules showed which variables have significant bearings for the outpatient visits and the findings revealed that those variables were examination flow, medical check-ups, queue length, doctor's expertise, and health treatment objectives.

The calculation process is repeated until all partitioned attributes. Once, it was completely calculated using the steps outlined above, a decision tree was drawn based on the predetermined node as shown in Figure 2 below.

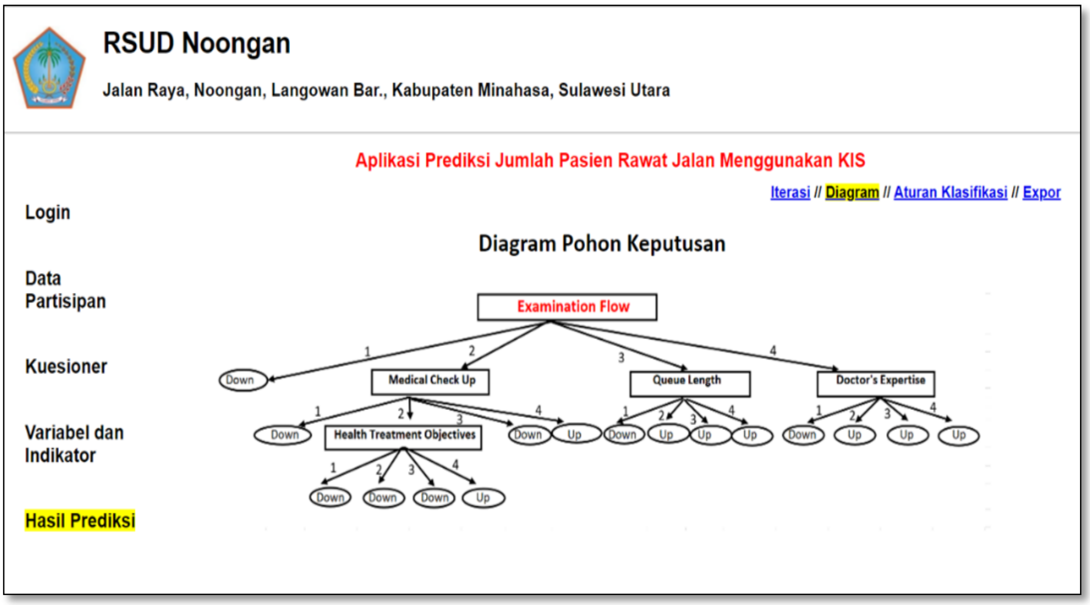


Figure 2 The Decision Tree Built by the C4.5 Algorithm

3.4 *Validation*

The dataset for the test data contained 48 responses from the participants. Using this dataset to test the C4.5 algorithm in terms of its accuracy, precision, recall, and F-measure to predict the outpatient visits using JKN-KIS at Noongan Hospital as illustrated in Table 6. The mean was calculated by summing up all the scores of the eight variables and then dividing the total score by eight. For the prediction value, if one data has a mean score below 2.5, was it considered being 'Down' in the number of outpatient visits. On the other hand, the prediction value was regarded as 'Up' in the number of outpatient visits when the mean score was above 2.5.

Table 6 The Test Dataset

No.	Total Score	Mean	Prediction Value	Actual Value (C4.5 Algorithm)	Classification Comparison
1	23	2.875	Up	Down	False
2	27	3.375	Up	Up	True
3	25	3.125	Up	Up	True
4	23	2.875	Up	Up	True
5	21	2.625	Up	Up	True
6	30	3.75	Up	Up	True
7	25	3.125	Up	Up	True
8	27	3.375	Up	Up	True
9	22	2.75	Up	Up	True
10	23	2.875	Up	Up	True
11	25	3.125	Up	Up	True
12	25	3.125	Up	Up	True
13	22	2.75	Up	Up	True
14	24	3	Up	Up	True
15	22	2.75	Up	Up	True
16	30	3.75	Up	Up	True
17	22	2.75	Up	Up	True
18	30	3.75	Up	Up	True
19	21	2.625	Up	Down	False
20	29	3.625	Up	Up	True
21	22	2.75	Up	Up	True

22	23	2.875	Up	Up	True
23	23	2.875	Up	Up	True
24	23	2.875	Up	Up	True
25	19	2.375	Down	Up	False
26	23	2.875	Up	Up	True
27	28	3.5	Up	Up	True
28	23	2.875	Up	Up	True
29	24	3	Up	Up	True
30	25	3.125	Up	Up	True
31	26	3.25	Up	Up	True
32	28	3.5	Up	Up	True
33	22	2.75	Up	Up	True
34	28	3.5	Up	Up	True
35	23	2.875	Up	Up	True
36	28	3.5	Up	Up	True
37	25	3.125	Up	Up	True
38	14	1.75	Down	Down	True
39	18	2.25	Down	Down	True
40	14	1.75	Down	Down	True
41	18	2.25	Down	Down	True
42	20	2.5	Down	Down	True
43	19	2.375	Down	Down	True
44	23	2.875	Up	Up	True
45	20	2.5	Down	Up	False
46	22	2.75	Up	Up	True
47	24	3	Up	Up	True
48	27	3.375	Up	Up	True

Based on the dataset in Table 5, the total number of outpatient visits using JKN-KIS at Noongan Hospital has increased by 83.33% since the pandemic of Covid-19. To further assess how well the C4.5 algorithm can classify the datasets using the eight variables and thus provide a prediction on the total number of outpatient visits. The number of prediction values and actual values in terms of true positive (TP), false positive (FP), false negative (FN), and true negative (TN) were 38, 2, 2, and 6 respectively. The findings revealed that the classification accuracy was 91,7% while the precision, recall, and F-Measure had the same results of 95% relevancy. These results indicated that the model's performance is considerably high due to its ability to classify the data and thus provide a reliable prediction of the number of outpatient visits using JKN-KIS at Noongan Hospital.

4. CONCLUSION

The following conclusions were drawn based on the research findings:

1. The C4.5 algorithm can accurately classify the data to the number of outpatient visits using JKN-KIS at Noongan Hospital. There were 8 variables used to predict the number of outpatient visits as the participants about a medical check-up, referral, extension of BPJS service, queue length, health treatment objectivities, doctor's expertise, examination flow, and hospital service satisfaction.
2. The model built by the C4.5 algorithm-generated 20 classification rules to predict the number of outpatient visits using JKN-KIS at Noongan Hospital.
3. The performance of a predictive model using the C4.5 algorithm for this research is considered high since the model has 91,7% of classification accuracy. This model also has a high precision (95%), indicating a low false-positive rate and a high recall (95%) related to a low false-negative rate.
4. The number of outpatient visits using JKN-NIS at this hospital has steadfastly increased by 83,33% since the pandemic of Covid-19.

5. Examination flow, medical check-ups, queue length, doctor's expertise, and health treatment objectives are the most influencing factors in the outpatient visits at Noongan Hospital. This hospital must find ways to improve the efficiency of those factors to increase its patient satisfaction.
6. This application can help the management of the hospital not only to allocate healthcare resources better but also to increase the number of outpatient visits by improving the efficiency of outpatient services.

5. FUTURE WORKS

There are several recommendations provided for future works as follows:

1. There is a need for more sample collection periodically in the longer term that can represent the population, be valid, and unusual, especially in data testing to produce an efficient and effective use in predicting the number of outpatients using JKN-KIS at Noongan Hospital during the Covid-19 pandemic Covid-19.
2. As a new variant of the mutation of the virus that causes Covid-19 appeared, it resulted in an increase in the number of positive cases and the number of in-patients. This also impacts how outpatient care is delivered at this hospital. Preventive health care practices are implemented to decrease the risk of transmitting the virus to either patients or health care workers. Hence, it is highly suggested to examine the data on changes in outpatient visit volume at this hospital after the new variant of Covid-19.
3. The Covid-19 vaccination is required for outpatient visits at many hospitals to reduce the risk of exposure. It is recommended to investigate the impact of Covid-19 vaccination on outpatient visits at Noongan Hospital so that the hospital management can make informed decisions.
4.

ACKNOWLEDGEMENT

We would like to thank our research participants who had spent time filling in the questionnaires. Also, we are grateful for the continuous support and motivation from Universitas Katolik De La Salle Manado.

REFERENCES

[1] BPJS. Kesehatan, 12 November 2020, Panduan Layanan Bagi Peserta Jaminan Kesehatan Nasional Kartu Indonesia Sehat (JKN-KIS), https://bpjs-kesehatan.go.id/bpjs/dmdocuments/e945442b347fb8dde4159327badc15b9.pdf, diakses tgl 1 November 2021.

[2] Budi, O, Fasilitas Kesehatan Tingkat Pertama (FKTP) dan Cakupannya, https://lifepal.co.id/media/fktp-adalah-fasilitas-kesehatan-tingkat-pertama, diakses tgl 10 November 2021.

[3] BPJS. Kesehatan, 3Maret 2017, Pahami Lebih Dalam tentang Sistem Rujukan Berjenjang dan Pola Pembayaran BPJS Kesehatan ke Faskes, https://www.bpjs-kesehatan.go.id/bpjs//unduh/index/269, diakses 10 November 2021.

[4] Haruna, A. A, Yahaya, B. Z, Oye, N. D, Muhammad, L. J, Garba, E. J, and Jung, L. T, An Improved C4.5 Data Mining Driven Algorithm for the Diagnosis, *2019 International Conference on Digitization (ICD)*, Sharjah, United Arab Emirates, 18-19 Maret 2019.

[5] Kantardzic, M, 2020, *Data Mining: Concepts, Models, Methods, and Algorithms*, John Wiley & Sons, Inc, New Jersey.

[6] Andriani, A, 2017. Klasifikasi Berbasis Algoritma C4.5 untuk Deteksi, *Seminar Nasional Informatika Medis,* Vol 8, Ed 1, 75-75.

[7] Noviandi, 2018, Implementasi Algoritma Decision Tree C4.5 Untuk Prediksi Penyakit Diabetes, *Indonesian of Health Information Management Journal,* No.1, Vol. 6, 1-5.

[8] Kasmad, Ahidin, U, and Wijayanti, K. D, 2018, Pengaruh Strategi Periklanan Terhadap Keputusan Mahasiswa(Studi Kasus Pada Stikes Pertamina Bina Medika Jakarta), *Jurnal Pemasaran Kompetitif,* No.2, Vol. 1, 20-31.

[9] Wikarsa, L and Kim, M-S, Automatic Generation Of Word-Emotion Lexicon For Multiple Sentiment Polarities On Social Media Texts, 2019 *ICIC Express Letters: An International Journal of Research and Surveys,* No. 4, Vol. 13, 317-324.

[10] Kurniawan, , Indriati, and Adinugroho, S, 2019, Analisis Sentimen Opini Film Menggunakan Metode Naïve Bayes dan Lexicon Based Features, *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer,* No. 9, Vol. 3, 8335-8342.

[11] Zahra, R. K and Rina, N, 2018, Pengaruh Celebrity Endorser Hamidah Rachmayanti Terhadap Keputusan Pembelian Produk Online Shop Mayoutfit Di Kota Bandung, *Jurnal Lontar,* No.1, Vol.6, 43-57.

[12] Budiastuti, D, and Bandur, A, 2018, *Validitas Dan Reliabilitas Penelitian*, Mitra Wacana Media, Jakarta.