BAYESIAN NETWORKS FOR PREDICTION CHRONIC KIDNEY DISEASE

Idza Runianti Darusman[1], Ednawati Rainarli[2]

[1,2]Department of Informatics, University Computer Indonesia
Jl. Dipati Ukur No. 112-116, Coblong, Bandung
E-mail: idzarunianti@gmail.com[1], ednawati.rainarli@email.unikom.ac.id[2]

ABSTRACT

In previous studies on the prediction of chronic kidney disease, Lambodar J. et al. have used the Naive Bayes method to predict chronic kidney disease. From this study, the results obtained 95% accuracy. After an interview with Dr. Eva M Hidayat found that there are several variables in the data of chronic kidney disease that have relationships with other variables. This is in contrast to the independent concept which is a requirement in the use of the Naive Bayes method. Bayesian Networks is used to solve this problem. The structure of Bayesian Networks was made according to the results of the interview with Dr. Eva M Hidayat and according to literature study. Parameter estimation in Bayesian Networks is using the maximum likelihood estimation (MLE). For validation using a 5-fold cross-validation method. After the dataset is tested, it was found that the variables blood urea, serum creatine, and pedal edema have an influence on the predictions of chronic kidney disease. Besides that, the average accuracy is 98% in the condition class in the dataset was balanced. This proves that Bayesian Networks have good prediction performance than Naive Bayes which have the average accuracy is 97%.

Keywords : Prediction, chronic kidney disease, Bayesian networks, Naive Bayes, Maximum Likelihood Estimation.

1. INTRODUCTION

Kidney is the organ that serves to maintain the composition of the blood by preventing the buildup of waste and fluid balance existing in the body [1], Maintaining kidney health is very important because the kidney is one organ is important in maintaining a healthy body. Disorders of the kidney there is a lot one of them is chronic kidney.

Chronic kidney disease (CKD) is damage to the kidneys which causes the kidneys can not remove toxins and waste products of blood, with a marked presence of protein in the urine and reduced glomerular filtration rate that lasts for more than 3 months [1] [2], According to Professor of Medicine Faculty of Medicine, University of Indonesia (FKUI), Parlindungan Siregar Kompas.com quoted from this disease known generally as already reached stage three to four [3], Delays in treatment and detection of this disease resulted in the prevalence of deaths from chronic kidney disease (CKD) is high so early detection is needed to reduce the prevalence rate of death due to CKD.

In a previous study, Lambodar Narendra Jena and Ku. Kamila has conducted studies to detect chronic kidney disease either by usingnaive Bayes, In the study found that the accuracy rate of 95% and the time required to create and test models for 0:02 seconds [4], However, after the interview with Dr. Eva M Hidayat, it was found that there are several variables that have a relationship with the other variables so that research using Naive Bayes can still be improved.

Naive Bayes said to be 'naive' because it is assumed that each feature is conditionally independent or it can be said that in the case remedy every feature has no relationship [5]. So naive Bayes can not resolve the case - specific cases, especially in cases where each feature requires correlation with each other. The method can be used to solve the problem of the relationship between these features, one using Bayesian networks.

In the study conducted by P.Fuster-Parra et al about the risk of heart disease by using Bayesian networks [6], From these studies it was found that the accuracy of the results obtained was 96%.

Therefore in this research will be a prediction of chronic kidney disease by using Bayesian networks seen from the features (predictor variables) that are already available.

2. THEORETICAL BASIS

2.1 Chronic Kidney Disease

Chronic kidney disease (CKD) is a kidney disease in which the kidneys have Glomerular filtration rate (GFR) below 60 mL / min / 1.73m2 for more than 3 months or evidence of kidney damage for more than 3 months [2],

In chronic kidney there are several symptoms - symptoms beginning to be felt by the patient, such as: high blood pressure, changes in urine, blood in the urine, loss of appetite, fatigue, nausea and

vomiting, insomnia, headaches [2], The factors of chronic kidney disease, namely [2] :

1. Hypertension
2. Diabetes mellitus
3. Family history of the disease CKD
4. Obesity
5. Age over 60 years

When a person has been suspected of having chronic kidney disease, the patient must go through several tests of kidney function as [2] :

1. The test for albumin (type of protein) and / or blood in the urine
2. A blood test to look at the levels of waste products in the blood and calculate glomerular filtration rate.
3. Blood pressure tests if kidney disease is caused by high blood pressure which can damage blood vessels in the kidneys.
4. CT scanto take pictures of the kidneys and urinary trace. This test for renal size and location of the tumor or kidney stones.

Based glomerulusnya filtration rate, chronic kidney disease is divided into several levels [2] which can be seen in Table 1

Table 1. Levels of Chronic Kidney Disease

	Level	Kidney condition
Early Stages	1	normal eGFR ≥ 90ml / min / 1,73m2
	2	EGFR decrease slightly between 60-89mL / min / 1,73m2
Middle Stages	3a	Mild-moderate decline in eGFR between 45-59mL / min / 1,73m2
	3b	Moderate-severe decline in eGFR between 30-44mL / min / 1,73m2
	4	The decline in heavy eGFR between 15-29 mL / min / 1,73m2
end Stages	5	Kidney failure, a decrease in eGFR <15m L / min / 1,73m2

2.2 Probability

Probability is the likelihood that an event [7], Probability theory has to do with different results [8], When the experiment is already defined, then the collection of all the results is called the sample space [8],

Mathematically, the sample space is the set while the result is called an element of the set [8], In the space limited samples, each sample chamber is called with the event and part of the sample chamber has an element called a basic events [8], The Rules of opportunities for events A, B, and are as follows;∅[9]

1. $0 \leq P(A) \leq 1$

   Probability of occurrence value lies between 0 and 1 inclusive.

2. $P(0) = \emptyset$

   Opportunities empty set is 0

3. $P(A) \leq= 1 - P(A)$

   Opportunities complement of an event is one - the odds of these events

4. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

2.2.1 Conditional probability

The conditional probability is the likelihood of an event which depends on other events [10], Conditional probability can use Bayes theorm as its decision-making.

Bayesian theorm This is a way to cope with uncertainty by using Bayes formula [11] which can be seen in equation 1

$$P(H|E) = \frac{P(E|H)\,P(H)}{P(E)} \qquad (1)$$

Where :

P (H | E): The probability of the hypothesis (H) if there is evidence

P (E | H): The probability of the emergence of evidence (E) if known hypothesis (H)

P (H): The probability of occurrence of the hypothesis (H) without seeing any evidence

P (E): The probability of the emergence of evidence (E)

2.3 Artificial Intelligence

According to the Oxford Dictionary Artficial Intelligence (AI) is the theory and development of computer systems to be able to do work that normally requires human intelligence, such as visual perception, speech recognition, decision makers, and translation between languages [12], The understanding of artificial intelligence, according to experts, namely [13] :

1. H.A.Simon (1987) defines "artificial intelligence (artificial intelligence) is an area of research, application, and instructions related to computer programming to do things in the human outlook is smart"
2. Rich and Knight (1991) defines "artificial intelligence (artificial intelligence) as a study of how to make computers do things - things that at this point can be done by humans"

2.4 Bayesian Networks

Bayesian Networks or commonly known as Belief Networks or probabilistic causal networks is a structural chart to represent the probabilities relationship between most of the variables by performing probabilistic inference to variable - the variable [8],

Bayesian networks has proven to be a powerful tool for finding the relationship between variables [6], Bayesian networks are widely used as decision makers in the field of banking [14], health [6], As well as decision makers on the use of transport [15],

In Bayesian networks, there are two learning undertaken, namely:

1. structure Learning

This learning stage is the stage of identifying the topology [6]models in Bayesian networks that will become a reference for decision-making. On learning structure there are three approaches: [16] :

a. Search-and-score-based method
b. Constraint-based approaches
c. Bayesian model of average

2. parameter Learning

This learning stage is the stage of calculating the estimate on any value in a variable or conditional probability in accordance with the existing topology [6], If the topology is not through the stages of learning the structure estimations can use two methods:

a. Maximum Likelihood Estimation (MLE)

MLE is a method often used to estimate the parameters of the model data is already known. A parameter value is said to be good views of how much is likely to produce the observed data [17], In chronic kidney disease predictive variables, there are two types of data, namely the binomial and multinomial. On the data then the binomial estimation calculation can be seen in equation 2 [17]

$$\theta = \frac{Na}{Na+Nb} \quad (2)$$

While on the variable data is multinomial, the calculation of the estimate can dilhat the equation 3 [17]

$$\theta = \frac{Nk}{\sum_{\lambda} N_{\lambda}} \quad (3)$$

In the variables that have a relationship with other variables, the calculation of the estimate can be seen in Equation 4 [17]

$$\theta_{xi|Pxi} = \frac{N(P_{xi} \cap Xi)}{Xi} \quad (4)$$

b. Maxiumum a Posteriori Estimation (MAP)

3. RESEARCH METHODS

The research method in this study is an experimental method because this method is a quantitative method where this method conducted an experiment to see results and to investigate the possibility of causation [18], The stages of this research method that can be seen in Figure 1

Picture 1. Research methods

3.1 Identification of Problems

Identification of the problem is the process of finding problems. In this study a problem that occurs is data chronic kidney disease [19] There are variables that have a relation to the other variables.

3.2 Data collection

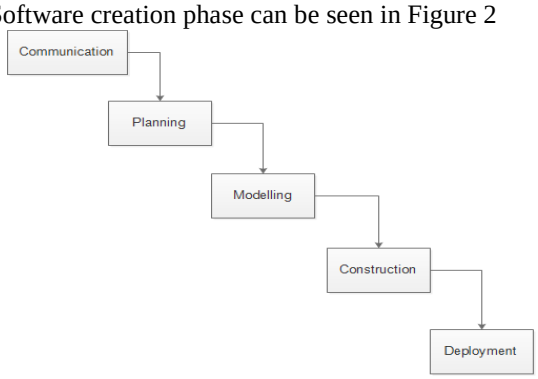
At this stage of data collection was done by conducting interviews with experts such as Dr. Eva M Hidayat. In addition, researchers also collect data by means of literature.

3.3 Data Analysis and Needs Algorithm

At this stage, the researchers conducted an analysis of the input data, datasets that was collected previously, as well as the output data. In addition, researchers also conducted understanding of the concept in Bayesian methods Networks.

3.4 Software Development

At this stage the researchers conducted a software development using Software Development Life Cycle (SDLC) using the waterfall method [13], Software creation phase can be seen in Figure 2



Picture 2. Software Development Phase

3.5 Testing

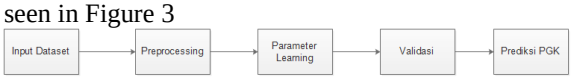
At this stage of testing by using a system that has been made on whether the predicted outcome of chronic kidney disease is good and true. The test is performed by a functional test on the system as well as testing ondatasets using k-fold cross validation and ROC curve.

3.6 Conclusion

At this stage the researchers draw conclusions with a view of the results of the accuracy of the prediction of chronic kidney disease.

4. RESULTS AND DISCUSSION

Chronic kidney disease prediction systems to be built consists of several main processes, namely preprocessing, parameter learning and validation. In general overview of the system to be built can be seen in Figure 3



Picture 3. System Overview

The input data (training data and test data) in chronic kidney disease prediction is obtained from chronic kidney dataset [19]which has been validated to Dr. Eva M Hidayat. The results of the validation shows that of the 25 variables which are only 13 variables used in which the 12 to predict chronic kidney disease and one variable is the class predictions. Variables - variables used are shown in Table 2

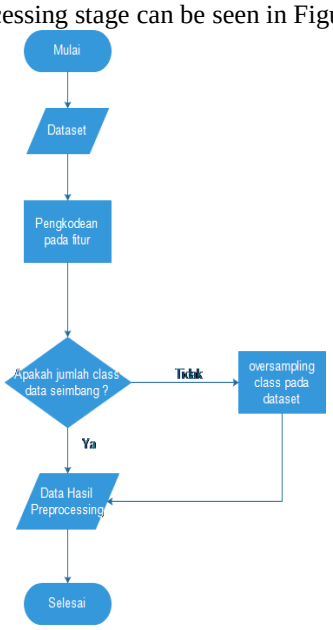
Table 2. Input Data Description

variables	Description	Type	Value
Age	Age	Numerical	Age in years

BP	Blood pressure	Numerical	mm / Hg
AL	Albumin	Nominal	{0,1,2, 3,4,5}
BGR	Random blood glucose	Numerical	Mgs / dl
BU	Blood urea	Numerical	Mgs / dl
SC	serum creatinine	Numerical	Mgs / dl
SOD	Sodium	Numerical	mEq / L
POT	potassium	Numerical	mEq / L
HTN	Hypertention	Nominal	Yes, No
DM	diabetic mellitus	Nominal	Yes, No
APPET	appetite	Nominal	Good, Poor
PE	pedal edema	Nominal	Yes, No
Class	Class	Nominal	Yes, No

4.1 Pre-processing

In the preprocessing stage of data will do the encoding and checked whether the classes in the data are balanced or not, if the data is not balanced then the class of data to be offset by using the technique of random oversampling which is a technique that is done by copying and repetition of existing data on the minority so Data equilibrium is reached or engineering sample copy the original data as repeated in large numbers [20, 21], The flow of the preprocessing stage can be seen in Figure 4



Picture 4, Preprocessing Flow

Nomial data encoding is 1 and 0. While data numerical coding is done by getting the results range from literature [22, 23, 24, 25, 26] and the age variable refers to the encoding of existing research-Parra P.Fuster [6], Encoding variables can be seen in Table 3

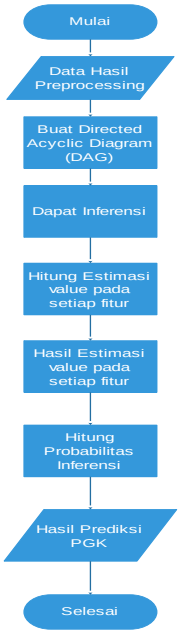
Table 3. encoding variable

variables	Value	Information
AL	1	There content of albumin in the urine (1-5)
	0	No content of albumin in the urine (0)
HTN	1	Suffer from hypertension (yes)
	0	Do not suffer from hypertension (no)
DM	1	Diabetes mellitus (yes)
	0	Do not suffer from diabetes mellitus (no)
APPET	1	Appetite is good (good)
	0	Appetite is not good (poor)
PE	1	Swelling in the legs (yes)
	0	Not having swelling in the legs (no)
Age	1	<18 years
	2	18-35 years
	3	35-55 years
	4	> 55 years
BP	1	Normal Blood Pressure (<85 mmHg)
	2	High Blood Pressure - Normal (85 mmHg - 89 mmHg)
	3	High Blood Pressure (≥ 90 mmHg)
SOD	1	Sodium Low (<135 mEq / L)
	2	Normal sodium (135 mEq / L - 145 mEq / L)
	3	High sodium (> 145 mEq / L)
POT	1	Low Potassium (≤ 3.5 mEq / L)
	2	Normal potassium (3.6 mEq / L - 5.2 mEq / L)
	3	High Potassium (≥ 5.3 mEq / L)
BGR	1	Normal Blood Sugar (<100 mgs / dl)
	0	High Blood Sugar (≥ 100 mgs / dl)
BU	1	Blood Urea Nitrogen normal (7 mg / dL - 20 mg / dL)
	0	Blood Urea Nitrogen high (≥ 20 mg / dL)
SC	1	Creatine serum normal (0.8 mg / dL - 1.2 mg / dL)
	0	Creatine serum High (> 1.2 mg / dL)

CLASS	1	Diagnosed with chronic kidney disease (ckd)
	0	Not diagnosed with chronic kidney disease (notckd)

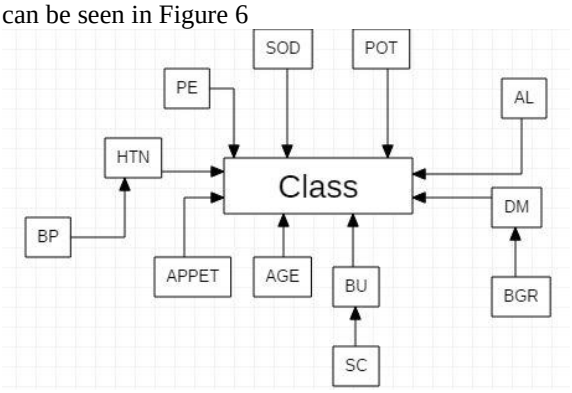
4.2 Learning Bayesian Networks

After the preprocessing stage, the next stage is the learning stage in the method of Bayesian networks. Learning phase can be seen in Figure 5



Picture 5. Learning Bayesian Networks Flow

At the stage of learning structure, topology has been made based on interviews and literature. Topology for prediction of chronic kidney disease can be seen in Figure 6



Picture 6. Topology for prediction Chronic Kidney Disease

Seen from Figure 5 the calculation of the decision or inference of chronic kidney disease can be seen in the equation 5

$$P\begin{pmatrix}\text{class, age, appet, pe, al, bu, sc, dm, bgr,}\\ \text{htn, pot, sod, bp}\end{pmatrix} = P(\text{age})P(\text{appet})P(\text{sod})P(\text{pot})P(\text{pe})P(\text{al})P(\text{htn}|\text{bp})P(\text{dm}|\text{bgr})P(\text{bu}|\text{sc}) \quad (5)$$
$$P(\text{class}|\text{htn, dm, sod, pot, appet, age, al, pe, bu})$$

The next stage is the stage of the calculation of estimated values for each variable determinant of chronic kidney disease prediction. At this stage of the method usedmaximum likelihood estimation (MLE). On the data in the form binomial variables, using equation 2, the estimation results obtained on PE variables that can be seen in Table 4

Table 4. Estimated Value In Binomial Data

PE	P (PE)
1	0.18
0	0.82

On the variable data in the form multinomial using equation 3, the estimation results obtained on SOD variables that can be seen in Table 5

Table 5. Results Estimates On Multinomial Data

SOD	P (SOD)
1	0.19
2	0.67
3	0.14

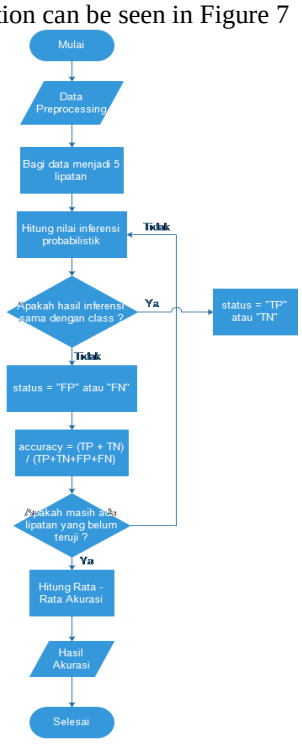
On the other variables that have a relationship then the calculation of estimated using equation 4. The results of the estimation of DM variables that have a relationship with BGR variables can be seen in Table 6

Table 6. Estimation Results In Related Variables

BGR	P (DM | BGR)	
	DM	
	1	0
1	0.14	0.86
0	0.36	0.64

4.3 Validation

Once the data is done by calculating the decision making probabilistic inference using the equation 5, the next step is validation by counting accuracy the method of k-fold cross validation and ROC curve. Flow validation can be seen in Figure 7



Picture 7. Validation Flow Prediction Chronic Kidney Disease

k-fold cross validation is a technique used to estimate the error when testing where this estimation can be used to view the best model and give an idea of testing the model chosen misunderstanding least the last [27],

Calculation accuracy by using ROC curve which is a technique of calculation accuracy by using a graph to visualize, organize, and selects a classification depending on performance [28], In the ROC curve created a matrix to compare the original data and the test data can be seen in Figure 8

	POST	NEG	
POST	TP	FP	
NEG	FN	TN	
	P	N	P + N

Picture 8. Matrix ROC curve

Tests on the ROC curve, there are three criteria: accuracy, sensitivity, and specificity that the calculation can be seen in equation 6 to Equation 8

$$\text{Accuracy} = \frac{TP+TN}{P+N} \quad (6)$$

$$\text{sensitivity} = \frac{TP}{P} \quad (7)$$

$$\text{specificity} = \frac{TN}{N} \quad (8)$$

4.4 Testing

In this testing phase dataset included by way of import. The system will perform preprocessing stage when data is entered so that data stored in the database is the result of data preprocessing. Once the data is entered, the next stage is the calculation of the estimate. The estimated value is in getting will be of value for the calculation of inference on the test data.

4.4.1 Accuracy Testing

Accuracy testing was conducted to determine the level of accuracy that is done by the system using a Bayesian algorithm. Testing is done using k-fold cross validation with the number k = 5 where the dataset will be divided into 5 parts: K1, K2, K3, K4, K5 to iteratively test 5 times.

The amount of data used was 260 with each - each of data (K) was 52 of data where the number of class predictions are equally Experimental testing was conducted five times, with each dataset will alternately be training data and test data. Scenario testing can be seen in Table 7

Table 7. Testing Scenario

No.	training	testing
1	K1, K3, K4, K5	K2
2	K1, K2, K4, K5	K3
3	K1, K2, K3, K5	K4
4	K1, K2, K3, K4	K5
5	K2, K3, K4, K5	K1

After calculation by using equations 5 and compared with the test data class the original data. By using the k-fold cross validation accuracy results obtained are shown in Table 8

Table 8. Accuracy Testing Results

No.	testing	accuracy
1	K2	0,981
2	K3	.942
3	K4	0,981
4	K5	1
5	K1	1
Average		0,981

In addition, by using the same test data validated by using ROC curve. By using equation 7 and equation 8, the test results can be seen in Table 9

Table 9. Testing Results ROC curve

No.	testing	sensitivity	specificity
1	K2	0,962	1
2	K3	0,885	1
3	K4	0,962	1
4	K5	1	1
5	K1	1	1
Average		0,962	1

Judging from the test results on the dataset by using the ROC curve, the area under the curve (AUC) in chronic kidney disease prediction testing can be seen in Figure 9

Picture 9. Results of ROC curve

5. CONCLUSIONS

After a test on the dataset by using k-fold cross validation found that the value - average accuracy is equal to 98.1%. It can be concluded that the predicted outcome of chronic kidney disease by using Bayesian networks increased compared with the predicted outcome of chronic kidney disease in the previous study using Naive Bayes method. In addition, the calculation accuracy by using the ROC curve when seen from the area under the curve model used to predict chronic kidney disease is not good because the area under the curve is unknown. This is because in a learning phase the construction of structures (structural learning) made sure.

The suggestion of this research is on stage structural learning should be built using existing approach on Bayesian networks such as search-score based, constraint-based methods, and Bayesian averaging models, in addition to the variables used to predict chronic kidney disease can be used all the variables to see other relationships between variables.

ACKNOWLEDGEMENT

Thanks to Dr. Eva M Hidayat as validator who has provided time, knowledge, and direction so I can finish this research.

BIBLIOGRAPHY

[1] InfoDatin, Situasi Penyakit Ginjal Kronis, Jakarta: Kementrian Kesehatan RI, 2017.

[2] K. H. Australia, "All About Chronic Kidney Disease (CKD)," Kidney Health Australia, Mellbourn, 2015.

[3] C. C. Amiranti, "Waspadai Gagal Ginjal The Silent Killer," Kompas, [Online]. Available: https://lifestyle.kompas.com/read/2017/02/27/070700623/waspadai.gagal.ginjal.the.silent.killer.. [Diakses 24 Maret 2018].

[4] L. Jena dan N. Kamila, "Distributed Data Mining Classification Algorithm for Prediction Chronic-Kidney-Disease," International Journal of Engineering Research in Management&Technology, vol. IV, no. 11, pp. 110-118, 2015.

[5] K. Collins-Thompson, "Naive Bayes," Coursera, [Online]. Available: https://www.coursera.org/learn/python-machine-learning/lecture/0XFms/naive-bayes-classifiers. [Diakses 10 Mei 2018].

[6] P.Fuster-Parra.et.al, "Bayesian network modelling : A case study of an epidemologic system analysis of cardiovascular risk," Comput. Methods Programs Biomed.

[7] M. Spiegel dan L. J. Stephens, Schaum's Outlines : Statistik Edisi Ketiga, Jakarta: Erlangga, 2007.

[8] R. E. Neapolitan, Learning Bayesian Networks, New Jersey: Pearson Prentice Hall, 2004.

[9] S. Nughroho, Dasar – Dasar Metode Statistika, Jakarta: Grasindo, 2007, pp. 6, 53-54.

[10] D. Afriyani, Matematika Kelompok Teknologi, Kesehatan, dan Pertanian, Bandung: Grafindo Media Pratama, 2008.

[11] P. S. Dewi, "SISTEM PAKAR DIAGNOSIS PENYAKIT IKAN KOI," Jurnal Ilmiah Komputer dan Informatika (KOMPUTA), vol. 4, no. 1, pp. 25-32, 2015.

[12] O. University, "English Oxford Living Dictionaries," Oxford University, [Online]. Available: https://en.oxforddictionaries.com/definition/artificial_intelligence. [Diakses 18 April 2018].

[13] T. Harihayati, "Sistem Pakar Mendiagnosa Penyakit Umum Yang Sering Diderita Balita Berbasis Web Di Dinas Kesehatan Kota Bandung," Jurnal Komputer dan Informatika (KOMPUTA), vol. 1, no. 1, pp. 65-69, 2012.

[14] M. W. Triki, "Bank Credit Risk: Evidence from Tunisia using Bayesian Networks," Journal of Accounting, Finance, and Auditing Studies, vol. III, no. 3, pp. 93 - 107, 2017.

[15] Q. Wang, H. Sun dan Q. Zhang, "A Bayesian Network Model on the Public Bicycle Choice," Hindawi, vol. 2017, pp. 1 - 13, 2017.

[16] C. Berzan, An Exploration of Structure Learning in Bayesian Networks, Medford: Tufts University, 2012.

[17] N. Friedman, Learning Bayesian Networks From Data, California: U.C Berkeley, 1998.

[18] Suryana, Metode Penelitian, Bandung: Universitas Pendidikan Indonesia, 2010.

[19] M. Lichman, "UCI Machine Learning Repository," 2013. [Online]. Available: http://archive.ics.uci.edu/ml. [Diakses 28 Februari 2017].

[20] T. Hoens dan N. V. Chawla, Imbalanced Datasets : From Sampling To Classifier, Notre Dame: John Wiley & Sons, Inc., 2013.

[21] Xionggao.et.al, "SROT: Sparse representation-based over-sampling technique for classification of imbalanced dataset," IOP Conference Series : Earth and Environmetal Science, vol. 81, no. 1, pp. 1-10, 2017.

[22] M. Clinic, "Blood Urea Nitrogen (BUN) test," Mayo Clinic, [Online]. Available: https://www.mayoclinic.org/tests-procedures/blood-urea-nitrogen/about/pac-20384821. [Diakses 27 Juni 2018].

[23] M. Clinic, "Creatine Test," Mayo Clinic, [Online]. Available: https://www.mayoclinic.org/tests-procedures/creatinine-test/about/pac-20384646. [Diakses 27 Juni 2018].

[24] M. C. Staff, "High Potassium (Hyperkalemia)," Mayo Clinic, [Online]. Available: https://www.mayoclinic.org/symptoms/hyperkalemia/basics/definition/sym-20050776. [Diakses 27 Juny 2018].

[25] M. C. Staff, "Hyponatermia," Mayo Clinic, [Online]. Available: https://www.mayoclinic.org/diseases-conditions/hyponatremia/symptoms-causes/syc-20373711. [Diakses 27 Juni 2018].

[26] M. C. Staff, "Low Potassium (Hypokalemia)," Mayo Clinic, [Online]. Available: https://www.mayoclinic.org/symptoms/low-potassium/basics/definition/sym-20050632. [Diakses 27 Juni 2018].

[27] M. A. L. Fernandez, Cross Validation, London: London of School Hygine & Tropical Medicine, 2015.

[28] T. Fawcett, “An introduction to ROC analysis,” Elsevier, vol. 27, no. 12, pp. 861-874, 2006.

[29] K/DOQI, Clinical Practice Guidlines For Chronis Kidney Disease : Evaluation, Classification, and Stratification, New York: National Kidney Foundation, Inc, 2002.