IOP Conf. Series: Materials Science and Engineering **434** (2018) 012047 doi:10.1088/1757-899X/434/1/012047

Comparison of Naive Bayes and K-Nearest Neighbor methods to predict divorce issues

M Irfan, W Uriawan*, O T Kurahman, M A Ramdhani and I A Dahlia

Department of Informatics, Sunan Gunung Djati Bandung State Islamic University, Jl. AH Nasution No. 105, Bandung, West Java, Indonesia

*wisnu_u@uinsgd.ac.id

Abstract. Divorce in Cimahi every year continues to increase, each month to reach an average of 800 divorce cases, from the case of 75 cases of 75% household divorce cases, while the rest of the other cases, such as marriage isbat and inheritance. Based on the problem there is a need for prediction to find out how much divorce in each month. One of the techniques used to find divorce is by doing data processing to predict the occurrence of a divorce that is by using data mining techniques such as Naive Bayes algorithm and K-Nearest Neighbor. This algorithm has a high degree of accuracy in predicting. The best level of accuracy between the two algorithms can be determined by comparison. Comparison of algorithm aims to get the algorithm that is considered the fastest and accurate to make a prediction of a problem. Result of comparison of Naive Bayes and K-Nearest Neighbor algorithm can be concluded that Naive Bayes algorithm yield 72,5% accuracy and K-Nearest Neigbor algorithm yield 57,5% accuracy.

1. Introduction

Marriage in Islamic view is something sacred, meaning worship to Allah, follow the sunnah of the Prophet and carried out on the basis of sincerity, responsibility, and follow the legal provisions to be heeded.

But can not be denied differences in marriage often lead to quarrel between husband and wife. The emergence of various problems in marriage, such as the environment, infidelity, child issues, economic problems, their age at marriage, it can shake a marriage. According to one online media, divorce in the region of Bandung Regency, Cimahi City, every year continues to increase. Based on data from the Cimahi Religious Courts Office that oversees the area, since 2014 then, the case of divorce handled on average rose 25 percent. Registrar of Cimahi Religious Court, Saefuloh said that every month on average, he received 800 cases. Of the 800 cases, 75 percent of them are household divorce cases. While the rest of other cases, such as marriage isbat, inheritance and others.

Based on the level of divorce that every year continues to increase, it is necessary to take the form of a decision of the case to find out how much divorce in each month. One of the data processing techniques that is suitable in helping the prediction of the occurrence of a divorce is by using data mining techniques [1]–[6]. Data mining techniques that are often used are Naive Bayes and K-Nearest Neighbor [7]–[11]. The Naive Bayes algorithm is a statistical classifier that can be used to predict the probability of membership of a class. While K-Nearest Neighbor finds the closest distance between the data to be evaluated with the previous case data [12]–[14]. Of the two methods need to be a comparison to find

out which method is best among them so that the prediction process will be more effective, and accurate [2][11][12].

This paper is organized into four sections. First, the introduction section that explains the background, problems, and objectives of the study. Second, the research methods section that describes the review methods used for comparison of methods. Third, result and discussions. The last section is conclusion and suggestion.

2. Research methods

Software development method used by the author as a framework in the study with this case study is a prototype model. Prototype is a method in system development that is suitable for the level of new research and small scope [15].

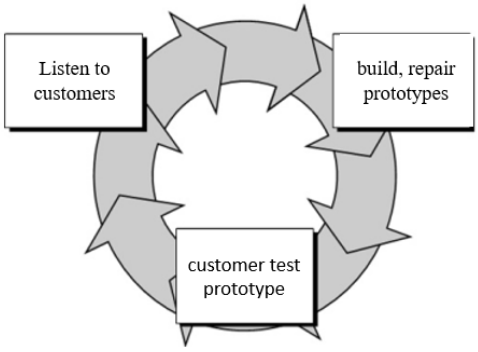

Figure 1. Prototype

2.1. Data mining

Simply, data mining refers to extracting or "mining" the knowledge of large amounts of data. Data mining is indeed one branch of computer science that is relatively new. And up to now people are still debating putting data mining in the field of science where data mining involves the integration of techniques from various disciplines. There are those who argue that data mining is nothing more than machine learning or statistical analysis that runs on a database. In scientific journals, data mining is also known as Knowledge Discovery in Databases (KDD) [1], [16], [17].

2.2. Naive bayes

The Naive Bayes Method is a statistical classification that can be used to predict the probability of membership of a class. Naive Bayes is based on the Bayes theorem that has similar classification capabilities to the decesion tree and neural network. Naive Bayes proved to have high accuracy and speed when applied into databases with large data [8], [18]–[20].

$$P(H|X) = \frac{\mathrm{P(X|H).P(H)}}{\mathrm{P(X)}} \quad (1)$$

Detail :

X	: Data with unknown class
H	: The hypothesis of X data is a class Specific
P (H | X)	: The probability of hypothesis H is based on condition X (posteriori probabilitas)
P (H)	: The probability of the hypothesis H (probability)
P (X | H)	Probability X based on condition at hypothesis H
P (X)	: The probability of X

2.3. K-nearest neighbor

The K-Nearest Neighbor method is a proximity search method between a new case and an old case based on a predetermined amount. Suppose necessary to find new patient by using solution from patient. To find the case of which patient to use, then the proximity of the new patient's case to the patient's old patient was calculated [8][9][10]. The case of the old patient with the greatest proximity would be considered a new patient problem.

$$\frac{\sum_{i=1}^{n} f(T_i, S_i)}{W_i} \quad (2)$$

Detail :

T : A new case
S : existing cases in storage
N : number of attributes
I : atriut individual between 1 - n
f : similarity function attributes i antarakasus T and case S
w : the weights assigned to the i-th attribute

2.4. Accuracy testing

Confusion matrix is a tool used to evaluate the classification model to estimate the correct or false object. A matrix of prediction that will be compared to the original class of input or in other words contains actual and predicted value information [8]. By calculation:

$$\text{Accuracy} = \frac{\text{the correct number of values}}{\text{total data}} \; x \; 100\%$$

3. Results and discussion

The object of this research is divorce data taken from Cimahi Religious Court, where the training data and data testing consist of several attributes such as plaintiff's age, age of the defendant, marriage age, child age and indicator. This will be a unity in the processing.

The workings of the Naive Bayes and K-nearest neighbor methods are almost identical, which distinguishes them during data groupings where the Naive Bayes method only categorizes the attribute data it uses, while the K-Nearest Neighbor method attributes are assigned a weighted value and the attribute data that has been grouped is assigned a value proximity to ease in the calculation process.

3.1 Example of calculation of naive bayes

In this research use case data of divorce case taken from Cimahi Religious Court year 2016 [9]. One example of case cases taken from Cimahi Religious Court as in the table below.

Table 1. Naive Bayes research data.

No	Age		Age of Marriage	Number of children	Indicator	Status
	Plaintiff	Defendant				
1	30	45	7	5	Murtad	Divorce
2	25	25	2	1	Berselisih	No
3	30	40	6	0	Moral	No
4	28	16	2	0	Berselisih	Divorce
5	19	19	1	1	Berselisih	No
6	20	19	2	1	Berselisih	No.
7	33	36	9	3	Menyakiti Jasmani	Divorce

3.2. Grouping attributes

This grouping aims to facilitate the calculation in the decision making process for prospective plaintiff before proceeding to the court case presented in the form of the following tables:

Table 2. Grouping age plaintiff.

No	Age Plaintiff
1	16-20
2	21-25
3	26-31
4	32-37
5	38-43
6	44-49
7	>50

Table 3. Grouping age defendant.

No	Age Defendant
1	16-20
2	21-25
3	26-31
4	32-37
5	38-43
6	44-49
7	>50

Table 4. Grouping age of marriage.

No	Age of Marriage
1	1-5
2	6-11
3	12-17
4	>18

Table 5. Grouping number of children.

No	Number of Children
1	0
2	1-4
3	>4

Table 6. Grouping indicator.

NO	Indicator
1	Meninggalkan Kewajiban
2	Berselisih
3	Menyakiti Jasmani
4	Moral
5	Kawin Dibawah Umur
6	Dihukum
7	Murtad
8	Cacat

3.3. New case

Table 7. Example of new case.

NO		1
Age	Plaintiff	19 th
	Defendant	25 th
Age of Marriage		2 th
Number of Children		1
Indikator		MORAL
STATUS		?

3.4. Calculation process cause of case

a. Find P (Y = DIVORCED) as the probability of a divorced person and P (Y = NO) as the probability of a person who is not divorced.

P (Y=DIVORCED) = 3 / 7 = 0,428

P (Y=NO) = 4 / 7 = 0,571

b. Counting the number of cases equal to the same class The amount of data P (X | Y = DIVORCED) with the description "Divorced" divided by the amount of divorced data and the amount of data P (X | Y = NO) with "NO" information divided by NO data amount.

P (USIA PENGGUGAT = 19 | Y = CERAI) = 0 / 3 = 0
P (USIA PENGGUGAT = 19 | Y = TIDAK) = 2 / 4 = 0,5
P (USIA TERGUGAT = 25 | Y = CERAI) = 0 / 3 = 0
P (USIA TERGUGAT = 25 | Y = TIDAK) = 1 / 4 = 0,25
P (LAMA PERKAWINAN = 2 | Y = CERAI) = 1 / 3 = 0,333
P (LAMA PERKAWINAN = 2 | Y = TIDAK) = 3 / 4 = 0,75
P (JUMLAH ANAK = 1 | Y = CERAI) = 1 / 3 = 0,333
P (JUMLAH ANAK = 1 | Y = TIDAK) = 2 / 4 = 0,5
P (INDIKATOR = MORAL | Y = CERAI) = 0 / 3 = 0
P (INDIKATOR = MORAL | Y = TIDAK) = 1 / 4 = 0,25

c. Multiply all results of the DIVORCED and NO variables

P(X|Y=CERAI) = P (USIA PENGGUGAT = 19 | Y = CERAI) x P (USIA TERGUGAT = 25 | Y = CERAI) x P (LAMA PERKAWINAN = 2 | Y = CERAI)x P (JUMLAH ANAK = 1 | Y = CERAI) x P (INDIKATOR = MORAL | Y = CERAI)
= 0 x 0 x 0,333 x 0,333 x 0
= 0

P(X|Y=TIDAK) = P (USIA PENGGUGAT = 19 | Y = TIDAK) x P (USIA TERGUGAT = 25 | Y = TIDAK) x P (LAMA PERKAWINAN = 2 | Y = TIDAK)xP (JUMLAH ANAK = 1 | Y = TIDAK) xP (INDIKATOR = MORAL | Y = TIDAK)
= 0,5 x 0,25 x 0,75 x 0,5 x 0,25
=0,011

P(X|Y=CERAI) x P (Y=CERAI) = 0 x 0,428
= 0

P(X|Y=TIDAK) x P (Y=TIDAK) = 0,011 x 0,571
= 0,006

d. Compare the results of the DIVORCED class and NOT Because the result (P | NO) is greater than (P | DIVORCED) then the decision is "NO"

3.5. Example of calculation of knn

Table 8. Example of new case.

NO		1
Age	Plaintiff	19 th
	Defendant	25 th
Age of Marriage		2 th
Number of Children		1
Indikator		MORAL
Status		?

Then from the existing case data sought the value of proximity from the age of the defendant, the age of the plaintiff, the length of marriage, the number of children, the indicator can be seen in the following table:

Table 9. KNN research data.

No	Age		Age of Marriage	Number of children	Indicator	Status
	Plaintiff	Defendant				
1	30	45	7	5	Murtad	Divorce
2	25	25	2	1	Berselisih	No
3	30	40	6	0	Moral	No
4	28	16	2	0	Berselisih	Divorce
5	19	19	1	1	Berselisih	No
6	20	19	2	1	Berselisih	No.
7	33	36	9	3	Menyakiti Jasmani	Divorce

3.6. Variable weighting

This weighting aims to facilitate the calculation in the decision-making process of old cases and new cases. The weight of this distance is given a value between 0 to 1. Value 0 means that if the attribute does not affect and vice versa 1 if the attribute is very influential.

Weight of one attribute with another attribute on the non-destination attribute can be defined with different values presented in the form of the following tables:

Table 10. Weighting.

No.	Variable	Weight
1	Age Plaintiff	0.6
2	Age Defendant	0.6
3	Age of Marriage	0.4
4	Number of Children	0.2
5	Indicator	0.8

3.7. Variable overview

The closeness between the values in the variable also needs to be defined. Here is the closeness between the values in the defined variables, namely:

3.7.1. Proximity of Plaintiff Age variable value. The following is the proximity of the Plaintiff Age variable values.

Table 11. Proximity of plaintiff age.

Age	16 - 20	21 - 25	26 - 31	32 - 37	38 – 43	44 - 49	> 49
16 - 20	1	0.777	0.666	0.555	0.444	0.333	0.111
21 - 25	0.777	1	0.857	0.714	0.571	0.428	0.142
26 - 31	0.666	0.857	1	0.8333	0.166	0.5	0.166
32 - 37	0.555	0.714	0.833	1	0.8	0.6	0.2
38 - 43	0.444	0.571	0.666	0.8	1	0.75	0.25
44 - 49	0.333	0.428	0.5	0.6	0.75	1	0.333
> 49	0.111	0.142	0.166	0.2	0.25	0.333	1

3.7.2. Proximity of the variable value of the Defendant's Age. The following model of proximity of the Defendant Age variable values is defined:

Table 12. Proximity of defendant age.

Age	16 - 20	21 - 25	26 - 31	32 - 37	38 - 43	44 - 49	> 49
16 - 20	1	0.777	0.666	0.555	0.444	0.333	0.111
21 - 25	0.777	1	0.857	0.714	0.571	0.428	0.142
26 - 31	0.666	0.857	1	0.8333	0.166	0.5	0.166
32 - 37	0.555	0.714	0.833	1	0.8	0.6	0.2
38 - 43	0.444	0.571	0.666	0.8	1	0.75	0.25
44 - 49	0.333	0.428	0.5	0.6	0.75	1	0.333
> 49	0.111	0.142	0.166	0.2	0.25	0.333	1

3.7.3. Proximity of married variable value value. The following is the proximity model of the variable value of the defined Marriage Length:

Table 13. Proximity of married age.

Age of Marriage	1 - 5	6 - 11	12 - 17	>17
1 – 15	1	0.833	0.333	0.166
6 – 11	0.8333	1	0.4	0.2
12 – 17	0.333	0.4	1	0.5
>17	0.166	0.2	0.5	1

3.7.4. Proximity value variable number of children. The following is the proximity model of the variable value of the Number of Children defined:

Table 14. Proximity of number of children.

Number of Children	0	1 - 4	>4
0	1	0.666	0.333
1 – 4	0.666	1	0.5
>4	0.333	0.5	1

3.7.5. Proximity of indicator variable values. The following is a proximity model of the Indicator variable value defined:

Table 15. Proximity of indicator.

INDIKATOR	Meninggalkan Kewajiban	Berselisih	Menyakiti Jasmani	Moral	Kawin dibawah Umur	Dihukum	Murtad	Cacat
Meninggalkan Kewajiban	1	0.9	0.8	0.7	0.6	0.5	0.4	0.3
Berselisih	0.9	1	0.889	0.778	0.667	0.555	0.444	0.333
Menyakiti Jasmani	0.8	0.889	1	0.875	0.75	0.625	0.5	0.375
Moral	0.7	0.778	0.875	1	0.875	0.714	0.571	0.428
Kawin dibawah Umur	0.6	0.667	0.75	0.857	1	0.833	0.667	0.5
Dihukum	0.5	0.556	0.625	0.714	0.833	1	0.8	0.6
Murtad	0.4	0.444	0.5	0.571	0.667	0.8	1	0.75
Cacat	0.3	0.333	0.375	0.428	0.5	0.6	0.75	1

3.8. Calculation process cause of case

To find out the resemblance between the old case and the new case. Before performing the calculation, to simplify and streamline the writing of the formula then do the definition of variables as follows:

NK = Value Proximity Case
NB = Weight Value
NK = Proximity of plaintiff's age new case against old case
NB = Weight of plaintiff's status
NK = Proximity of the defendant's age new case against old case
NK = Weight of defendant's age status
NB = Proximity of marital status of marriage new case against old case
NK = Weight of marriage duration
NB = Proximity status of the number of children new case against old case
NK = Weight of child number status
NB = Proximity status indicator case new to old case
NK = The indicator status weight
S = Distance

Here is the calculation process to find its similarity:

K1 Case

Proximity of new plaintiff's case status to old case = 0.666
Weight of plaintiff's age = 0.6
The proximity of the defendant's new age status to the old case = 0.399
Weight of defendant age = 0.6
Proximity of old status of new case marriage to old case = 0.833
Married status = 0.4
The proximity of the new child case status to the old case = 0.5
The status weight of the child count = 0.2
Proximity status of new case indicator to old case = 0.571
Status Weight indicator = 0.8

$$= \frac{(0.666x0.6) + (0.399x0.6) + (0.833x0.4) + (0.5x0.2) + (0.571x08)}{0.6 + 0.6 + 0.4 + 0.2 + 0.8}$$

$$= \frac{1.544}{2.6}$$

$$= 0.593$$

The above calculation is valid until case 7, so the results obtained, can be seen in the following table:

Table 16. The calculation results K1-K7.

No	Case	Proximity Age Plaintiff		Age Defendant		Age of Marriage		Number of Children		Indicator		Similarity
		NK	NB	NK	NB	NK	NB	NK	NB	NK	NB	
1	K1	0.666	0.6	0.333	0.6	0.833	0.4	0.5	0.2	0.571	0.8	0.593
2	K2	0.777	0.6	1	0.6	1	0.4	1	0.2	0.778	0.8	0.88
3	K3	0.666	0.6	0.571	0.6	0.833	0.4	0.666	0.2	1	0.8	0.771
4	K4	0.666	0.6	0.777	0.6	1	0.4	0.666	0.2	0.778	0.8	0.778
5	K5	1	0.6	0.777	0.6	1	0.4	1	0.2	0.778	0.8	0.802
6	K6	1	0.6	0.777	0.6	1	0.4	0.666	0.2	0.778	0.8	0.854
7	K7	0.555	0.6	0.714	0.6	0.883	0.4	1	0.2	0.875	0.8	0.766

3.9. Accuracy comparison

In the test of accuracy using two experiments with 20 different data testing, 130 training data are the same. The data will be tested to determine the accuracy generated by Naive Bayes and K-Nearest Neighbor algorithm, then will be analyzed so that can be known the number of true and false when doing the prediction on the data testing can be seen in the following table:

Table 17. Test of accuracy calculation 1.

	Naive Bayes	K-Nearest Neighbor
The amount of data tested	20	20
The data is correct	14	11
The data is wrong	6	9
Accuracy	70%	55%

Table 18. Test of accuracy calculation 2.

	Naive Bayes	K-Nearest Neighbor
The amount of data tested	20	20
The data is correct	15	12
The data is wrong	5	8
Accuracy	75%	65%

From two trials comparing the Naive Bayes and K-Nearest Neighbor algorithms. To predict divorce in Cimahi Religious Court with 20 different data testing resulted different percentage of accuracy also.

3.10. Results

Of the two accuracies above each algorithm has an average of 72.5% for the Naive Bayes algorithm and 57.5% for the K-Nearest Neighbor algorithm.

4. Conclusions and suggestions

4.1. Conclusions

Based on the research that has been done above, it can be concluded as follows:

a. The performance of Naive Bayes algorithm when making predictions does not take long because it has a high speed when applied to the database with large data and manual calculations are not complicated, while the K-Nearest Neighbor algorithm takes a long time because the distance calculation required from each new case in all the old cases and manual calculations quite complicated.
b. The Naive Bayes and K-Nearest Neighbor algorithms for divorce prediction in Cimahi Religious Court resulted in 72.5% accuracy for the niave bayes algorithm and 57.5% for the K-Nearest Neighbor algorithm using 20 data testing and 130 training data, the algorithm Naive Bayes is right to make predictions of divorce at the Cimahi Religious Court.

4.2. Suggestions

To improve performance and refine the research that has been made, suggestions for the development of similar systems can be compared with other algorithms such as Decesion Tree, Neural Network or Fuzzy in order to better understand algorithms in the process of predicting divorce in the Cimahi Religious Court [10].

References

[1] T Pang Ning 2006 Introduction to Data Mining (Boston: Pearson Addison Wesley)
[2] D Sa'Adillah Maylawati, M Irfan and W Budiawan Zulfikar 2017 Comparison between BIDE, PrefixSpan, and TRuleGrowth for Mining of Indonesian Text *J. Phys. Conf. Ser.* **801**(1)
[3] R Fauzan, Y Indrasary and N Muthia 2017 Sistem Pendukung Keputusan Penerimaan Beasiswa Bidik Misi di POLIBAN dengan Metode SAW Berbasis Web *JOIN (Jurnal Online Inform.)* **2**(2) pp. 79–83

[4] M A Ramdhani 2013 Metodologi Penelitian untuk Riset Teknologi Informasi (Bandung: UIN Sunan Gunung Djati Bandung)
[5] D S Maylawati, A Rahman, M I N Saputra, W Darmalaksana and M A Ramdhani 2018 Model of Citation Network Analysis using Sequence of Words as Structured Text Representation *IOP Conf. Ser. Mater. Sci. Eng.* **288**(1) p. 12048
[6] C Slamet, A R Atmadja, D S Maylawati, R S Lestari, W Darmalaksana and M A Ramdhani 2018 Automated Text Summarization for Indonesian Article Using Vector Space Mode *IOP Conf. Ser. Mater. Sci. Eng.* **288**(1) p. 12037
[7] D Syahid, Jumadi and D Nursantika 2016 Sistem Klasifikasi Jenis Tanaman Hias Daun Philodendron Menggunakan Metode K-Nearest Neighboor (Knn) Berdasarkan Nilai Hue, Saturation, Value (Hsv) *J. Online Inform.* **1**(1) pp. 20–23
[8] P Mata 2016 Perbandingan naive bayes classifier dengan nearest neighbor untuk identifikasi *JOIN (Jurnal Online Inform.)* **1**(2) pp. 82–86
[9] M A Naufal 2017 *Implementasi Metode Klasifikasi K-Nearest Neighbor (K-NN) Untuk Pengenalan Pola Batik Motif Lampung* (Lampung: Universitas Lampung)
[10] W B Zulfikar and N Lukman 2016 Perbandingan Naive Bayes Calssifier dengan Nearest Neighbor untuk Identifikasi Penyakit Mata in *JOIN (Jurnal Online Informatika)* **1**(2) pp. 82–86
[11] Y A Gerhana, W B Zulfikar, A H Ramdani and M A Ramdhani 2018 Implementation of Nearest Neighbor using HSV to Identify Skin Disease *IOP Conf. Ser. Mater. Sci. Eng.* **288**(1) p. 012153
[12] I Septiana, M Irfan and A R Atmadja 2016 Sistem Pendukung Keputusan Penentu Dosen Penguji Dan Pembimbing Tugas Akhir Menggunakan Fuzzy Multiple Attribute Decision Makingdengan Simple Additive Weighting (Studi Kasus: Jurusan Teknik Informatika Uin Sgd Bandung) *J. Online Inform.* **1**(1) pp. 43–50
[13] M Irfan 2015 Analisa pola asosiasi jalur masuk terhadap kelulusan mahasiswa dengan menggunakan metode fold-growth (studi kasus fakultas sains dan teknologi) *ISTEK* **9**(2)
[14] M Irfan, I Z Mutaqin and R G Utomo 2016 Implementation of Dynamic Time Warping algorithm on an Android based application to write and pronounce Hijaiyah letters *Proc. 2016 4th Int. Conf. Cyber IT Serv. Manag. CITSM 2016*
[15] G D Garson and M Khosrowpour 2008 Handbook of research on public information technology
[16] E Elisa 2017 Analisa dan Penerapan Algoritma C4 . 5 Dalam Data Mining Untuk Mengidentifikasi Faktor-Faktor Penyebab Kecelakaan Kerja Kontruksi PT . Arupadhatu Adisesanti *JOIN (Jurnal Online Inform.*) **2**(1) pp. 36–41
[17] J Laras Purwati Ayuningtias and Mohamad Irfan 2017 Analisa Perbandingan Logic Fuzzy Metode Tsukamoto, Sugeno, dan Mamdani (Studi Kasus : Prediksi Jumlah Pendaftar Mahasiswa Baru Fakultas Sains dan Teknologi Universitas Islam Negeri Sunan Gunung Djati Bandung) *J. Tek. Inform.* **10**(1) p. 19
[18] W B Zulfikar, M.Irfan, C N Alam and M Indra 2017 The comparation of text mining with Naive Bayes classifier, nearest neighbor, and decision tree to detect Indonesian swear words on Twitter in *2017 5th International Conference on Cyber and IT Service Management, CITSM 2017*
[19] W B Zulfikar, M Irfan, C N Alam and M. Indra 2017 The comparation of text mining with Naive Bayes classifier, nearest neighbor, and decision tree to detect Indonesian swear words on Twitter *2017 5th Int. Conf. Cyber IT Serv. Manag. CITSM 2017*
[20] Eska sarti Kundari 2015 Perbandingan kinerja metode Naive Bayes dan C4.5 dalam pengklasifikasian penyakit diabetes melitus di rumah sakit Kumala Siwi Kudus Universitas Dian Nuswantoro Semarang
[21] Salamun 2017 Penerapan Algoritma Nearest Neighbor dan CBR pada Expert System Penyimpangan Perilaku Seksual *JOIN (Jurnal Online Inform.*) **2**(2) pp. 63–70
[22] O Nurdiana, Jumadi and D Nursantika 2016 Perbandingan Metode Cosine Similarity Dengan

Metode Jaccard Similarity Pada Aplikasi Pencarian Terjemah Al-Qur'an Dalam Bahasa Indonesia *J. Online Inform.* **1**(1) pp. 59–63