Extraction Information on Incoming Mail Documents With Fuzzy K-Nearest Neighbor Algorithm (Fuzzy K-NN)

Adam Sulaiman[1], Nelly Indriani[2]

Informatics Engineering - Indonesian Computer University
Jl. Dipati Ukur 114 Bandung
E-mail: sulaimanadam@email.unikom.ac.id[1], nelly.indriani@email.unikom.ac.id[2]

ABSTRACT

Information extraction is one of the branches of text mining that aims to retrieve information from unstructured text into structured data. Document information extraction using rule-based classification has a problem of reducing accuracy in the characteristics of documents that are not identical. Therefore, this study uses a Fuzzy K-Nearest Neighbor (Fuzzy K-NN) learning machine as a document information extraction classification. The document used in this study is an incoming letter. The purpose of this study was to analyze how the level of performance of the Fuzzy K-NN classification in the extraction of incoming mail information that has the characteristics of the data is not identical and compares the level of accuracy with the K-NN classification. This study uses converting, filtering, tagging and tokenization processes, and uses boolean weighting as a weighting method. The feature extraction used has 11 parameters. Tests carried out using black box testing and calculation of accuracy with confusion matrix. The results of the accuracy testing showed that the accuracy of the Fuzzy K-NN classification was better than the K-NN classification with 87.52% compared to 85.17%. Based on these results, it can be concluded that the Fuzzy K-NN algorithm can be applied to a classification in the extraction of incoming mail document information.

Keywords: Information Extraction, Machine Learning, Fuzzy Logic, K-Nearest Neighbor, Incoming Letter.

1. INTRODUCTION

1.1 Background

Information extraction is a branch of the scope of text mining that aims to transform the results of the text mining process into the same root as the world of structured data in data mining [1] in which there is a classification method as a determinant of the data class [2] In this study information extraction is applied to incoming mail documents. In this study the meaning of an incoming letter is all official or official letters received by agencies or individuals [3].

Previous research on information extraction was carried out by Agny Ismaya. In this study a tool with rule-based classification was made to extract information on the Audit Result Report (LHP) document on the Local Government Financial Report (LKPD) [4]. From the results of the study it is known that the classification results on the test data have decreased. This happens because of the use of the wordmatch method (word match), so that if there is a slight difference in the document there will be an error in the classification. This problem may be solved if the classification method can determine the pattern of documents in terms of learning the similarity between documents or machine learning.

The documents used in the extraction of information in this study are incoming mail documents. Hermawan S. conducts administrative research in Kel. Jeruk Kab Sragen, by making an application that aims to handle the correspondence process [5]. However, in the results of the research the data storage of incoming mail is still entered manually so that the process is still the same as recording the data of the letter into the agenda book or typing manually into the data of the computer archive. By extracting information on incoming mail documents, information retrieval can be done automatically so that it will summarize the data input process manually. However, the idea of information-based extraction using rule-based is less able to handle the characteristics of data that are not identical, such as incoming letters, so that other methods are still needed to be able to handle the characteristic classification of data that is not identical such as incoming mail documents.

The method that will be used for the classification process is Fuzzy K-Nearest Neighbor. In

a study conducted by Zhang et al. [6] and in a study conducted by Satria N. et al. [7], it was said that the advantages of the fuzzy K-NN method lies in the accuracy level produced better than other classification methods such as Support Vector Machine (SVM), and K-Nearest Neighbor (K-NN). However, Fuzzy K-NN has not been applied in the extraction of document information, so it is unknown whether if applied the accuracy level of Fuzzy K-NN remains good or not.

Based on the previous exposure, then in this study used the method of Fuzzy K-Neurest Neighbor as a classification in information extraction. This is intended to process incoming mail documents into structured information and measure the level of accuracy of the Fuzzy K-NN algorithm in the extraction of information on incoming mail documents.

2. CONTENT OF RESEARCH

2.1 System Analysis

2.1.1 Process Description

In the system design will be built information extraction in incoming mail documents using Fuzzy K-NN classification .. Before weighting and classification, the input data is converting using the pdftables API, where this is done so that the reading of the text becomes easier and then through the preprocessing stage, namely filtering, tagging, and tokenization. After weighting and classifying test data using Fuzzy K-NN extraction of information and testing of classification results is carried out. The description of the process flow can be seen in Figure 1.

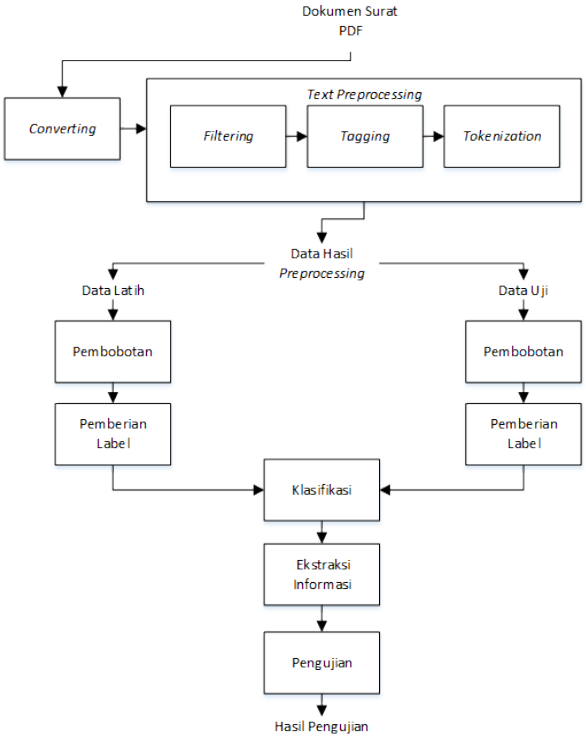
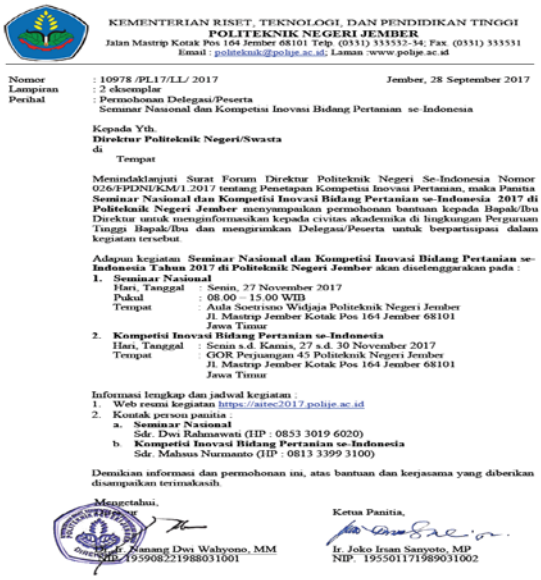


Figure 1. Process Description

2.1.2 Converting

Conversion of pdf to xml is the process of converting a letter file in pdf format to xml. This is done so that the text of the letter can be more easily read by the system. This conversion process is carried out by tools from the API that are on the domain pdftables.com. As an illustration of the conversion of pdf to xml is shown in Figure 2. and Figure 3.



KEMENTERIAN RISET, TEKNOLOGI, DAN PENDIDIKAN TINGGI
POLITEKNIK NEGERI JEMBER
Jalan Mastrip Kotak Pos 164 Jember 68101 Telp. (0331) 333532-34; Fax. (0331) 333531
Email : politeknik@polije.ac.id; Laman :www.polije.ac.id

Nomor : 10978 /PL17/LL/ 2017 Jember, 28 September 2017
Lampiran : 2 eksemplar
Perihal : Permohonan Delegasi/Peserta
Seminar Nasional dan Kompetisi Inovasi Bidang Pertanian se-Indonesia

Kepada Yth.
Direktur Politeknik Negeri/Swasta
di
Tempat

Menindaklanjuti Surat Forum Direktur Politeknik Negeri Se-Indonesia Nomor 026/FPDNI/KM/1.2017 tentang Penetapan Kompetisi Inovasi Pertanian, maka Panitia Seminar Nasional dan Kompetisi Inovasi Bidang Pertanian se-Indonesia 2017 di Politeknik Negeri Jember menyampaikan permohonan bantuan kepada Bapak/Ibu Direktur untuk menginformasikan kepada civitas akademika di lingkungan Perguruan Tinggi Bapak/Ibu dan mengirimkan Delegasi/Peserta untuk berpartisipasi dalam kegiatan tersebut.

Adapun kegiatan Seminar Nasional dan Kompetisi Inovasi Bidang Pertanian se-Indonesia Tahun 2017 di Politeknik Negeri Jember akan diselenggarakan pada :
1. Seminar Nasional
Hari, Tanggal : Senin, 27 November 2017
Pukul : 08.00 – 15.00 WIB
Tempat : Aula Soetrisno Widjaja Politeknik Negeri Jember
Jl. Mastrip Jember Kotak Pos 164 Jember 68101
Jawa Timur
2. Kompetisi Inovasi Bidang Pertanian se-Indonesia
Hari, Tanggal : Senin s.d. Kamis, 27 s.d. 30 November 2017
Tempat : GOR Penjaringan 45 Politeknik Negeri Jember
Jl. Mastrip Jember Kotak Pos 164 Jember 68101
Jawa Timur

Informasi lengkap dan jadwal kegiatan :
1. Web resmi kegiatan https://aitec2017.polije.ac.id
2. Kontak person panitia :
a. Seminar Nasional
Sdr. Dwi Rahmawati (HP : 0853 3019 6020)
b. Kompetisi Inovasi Bidang Pertanian se-Indonesia
Sdr. Mahsus Nurmanto (HP : 0813 3399 3100)

Demikian informasi dan permohonan ini, atas bantuan dan kerjasama yang diberikan disampaikan terimakasih.

Mengetahui,
Direktur
Ir. Nanang Dwi Wahyono, MM
NIP. 195908221988031001

Ketua Panitia,
Ir. Joko Irsan Sanyoto, MP
NIP. 195501171989031002

Figure 2. Input file Letters

```
<data-page="1" data-table="1"><tr><td></td><td></td><td colspan="9">KEMENTERIAN RISET, TEKNOLOGI, DAN PENDIDIKAN TINGGI</td></tr><tr><td></td><td></td><td></td><td></td><td colspan="6">POLITEKNIK NEGERI JEMBER</td><td></td></tr><tr><td></td><td colspan="10">Jalan Mastrip Kotak Pos 164 Jember 68101 Telp. (0331) 333532-34; Fax. (0331) 333531</td></tr><tr><td></td><td></td><td></td><td colspan="7">Email : politeknik@polije.ac.id; Laman :www.polije.ac.id</td><td></td></tr><tr><td>Nomor</td><td colspan="5">: 10978 /PL17/LL/ 2017</td><td></td><td></td><td colspan="3">Jember, 28 September 2017</td></tr><tr><td>Lampiran</td><td colspan="2">: 2 eksemplar</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr><tr><td>Perihal</td><td colspan="5">: Permohonan
```


Figure 3. Conversion Result Letter (.xml)

2.1.3 Text Preprocessing

Text Preprocessing is the stage of preparing an xml file into data that is ready for classification, extraction and testing. Text preprocessing consists of filtering, tagging, and tokenization.

a. Filtering

Results from pdf to xml conversion provide data that is still not suitable, many html tags are not needed. Therefore, it is necessary to do the conversion results filtering to clean unneeded tags. The following steps are steps for cleaning thesis data in Figure 4 and filtering rules in Table 1.

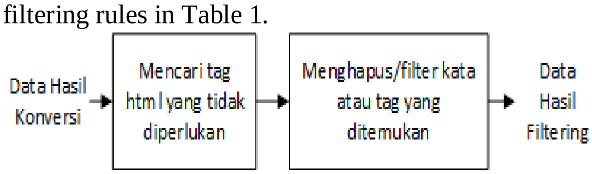


Figure 4. Stages of Filtering Processes

Table 1. Filtering Rules for

Signs / Words	**Action**
p, td, tr, </td>, </tr>	Dihapus
colspan="2", …, colspan="12"	
data-page="1"data-table="1"	
style="text-align: right	
Selain Tanda/Kata Diatas	Tidak ada aksi

Example results from the filtering process can be seen in Figure 5.

<><>KEMENTERIAN RISET, TEKNOLOGI, DAN PENDIDIKAN TINGGI<><><><><><><>POLITEKNIK NEGERI JEMBER<><><><>Jalan Masip Kotak Pos 164 Jember 68101 Telp. (0331) 333532-34; Fax. (0331) 333531<><><><><><>Email : politeknik@polije.ac.id; Laman :www.polije.ac.id<>Nomor<>: 10978 /PL17/LL/ 2017<><><>Jember, 28 September 2017<><>Lampiran<>:2 eksemplar<><><><><><><><><><><><>Perihal<>: Permohonan

<> < EDU POLI 6810 Emai <>: 1 2017 <> <

Picture 5 Example of Filtering Results

b) Tagging

After cleaning the tag, the next is the tagging process, the tagging process is done to tag the keyword and replace the tag into a delimiter which will be used in the tokenization process. The process steps can be seen in Figure 6 and the rules after being tagged in the keywords in Table 2.

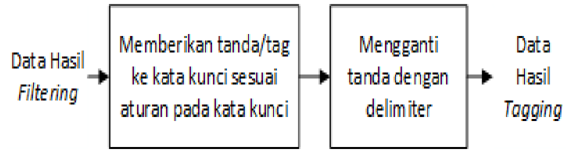


Figure 6. Stages of Tagging Processes

Table 2. Rules for Tagging

Condition	Action
Keadaan	Aksi
Memiliki *tag* <>, <> AWAL, dan AKHIR <>	Diganti delimiter |

Examples of the results of the filtering process can be seen in Figure 7.

KEMENTERIAN RISET, TEKNOLOGI, DAN PENDIDIKAN TINGGI POLITEKNIK NEGERI JEMBER|Jalan Masip Kotak Pos 164 Jember 68101 Telp. (0331) 333532-34; Fax. (0331) 333531 Email : politeknik@polije.ac.id; Laman :www.polije.ac.id |Nomor : 10978 /PL17/LL/ 2017 |Jember, 28 September 2017|Lampiran : 2 eksemplar|Perihal: Permohonan

Image 7. Example of Tagging Results

c. Tokenization

After the text has delimiter, the next process is to break the string into sentences according to delimiter. The process can be seen in Figure 8.

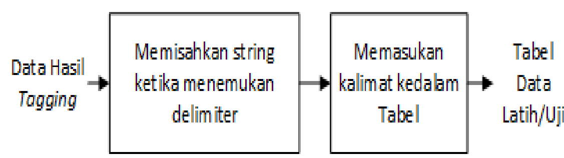


Figure 8. Stages of the Tokenization Process

Examples of the results of the filtering process can be seen in Table 3.

Table 3. Examples of Tokenization Results

No	Kalimat
1	KEMENTERIAN RISET, TEKNOLOGI, DAN PENDIDIKAN TINGGI POLITEKNIK NEGERI JEMBER
2	Jalan Masip Kotak Pos 164 Jember 68101 Telp. (0331) 333532-34; Fax. (0331) 333531 Email : politeknik@polije.ac.id; Laman :www.polije.ac.id
3	Nomor : 10978 /PL17/LL/ 2017
4	Jember, 28 September 2017
5	Lampiran: 2 eksemplar
6	Perihal: Permohonan

2.1.4 Weighting and Labeling

Weighting using a boolean weighting with 11 features used can be seen in Table 4. Value 1 if there is a feature and is 0 if not [8].

Table 4. Used Features

FEATURES		
X1	EMAIL	Have email / web
X2	ALLCAPS	All capital letters
X3	DIGIT	Have numbers
X4	CONTAINDASH	Have a line mark
X5	CONTAINSLASH	Have a slash
X6	INITIAL_KEYWORD	Have keywords
X7	LOCATION	Have anplace identity
X8	CONTAINCOLON	Have a colon
X9	STRING_LENGTH	Long Character> 250
X10	CONTAINCOMMA	Having a comma
X11	DATE	Having an identity date

After the next weighting is the labeling process. The training data label is used in the classification process while the test data label is used in the accuracy testing process with confusion matrix. The class label used can be seen in Table 5.

Table 5. Labels Used

No.	CLASS	LABEL CLASS
1.	1	1. Name of Instancy
2.	2	2. Address of Instancy
3.	3	3. Date of Letter
4.	4	4. Letter Number
5.	5	5. Appendix
6.	6	6. Subject
7.	7	7. Objective
8 .	8	8. Letter Contents

2.1.5 Fuzzy K-NN classification Analysis

classification phase is done after the process of weighing and labeling of training data is completed, it will be the classification of test data using the Fuzzy K-NN. The following is an explanation of the stages in the Fuzzy K-NN algorithm [9] for each test data:

1. Determine the value of k
2. Calculate the distance of the test and train document using the euclidean distance (1) method

$$\|x - x_j\| = \sqrt{\sum_{b=1}^{p} (x_b - x_{j,b})^2} \qquad (1)$$

Description:

$\|x - x_j\|$ = distance *euclidean* x test data with jth training data

p = Value max. from the calculated weight index

b = calculated weight index

x_b = b-weight index value in test data x

$x_{j,b}$ = b-weight index value in training data x_j

example calculates *euclidean distance distance* :

calculate distance of test data h1 with training data x_1

$$\|h1 - d_1\| = \sqrt{(h1_1 - d_{1,1})^2 + (h1_2 - d_{1,2})^2 + \cdots + (h1_{11} - d_{1,11})^2}$$

$$\|h1 - d_1\| = \sqrt{(0-0)^2 + (1-1)^2 + \cdots + (0-0)^2}$$

$$\|h1 - d_1\| = \sqrt{1} = 1$$

3. Take the neighbor's distance nearest number of k
4. The difference between Fuzzy K-NN from K-NN is the calculation of the membership value of each class with equation (2).

$$\mu_i(x) = \frac{\sum_{j=1}^{k} \mu_{ij} \left(\frac{1}{\|x - x_j\|^{\frac{2}{m-1}}}\right)}{\sum_{j=1}^{k} \left(1/\|x - x_j\|^{2/(m-1)}\right)} \qquad (2)$$

Description:

$\mu_i(x)$ = i-class membership value in x test data.

μ_{ij} = Neighbor membership value k (from training data □□) in class i (value 1 if neighbor class k from training data □□ equals class i, and value 0 if not) [9]

j = index training data

k = The amount of closest neighboring value taken

m = weight of the weight that is proportional to the distance between x and x_j.

Class membership value with fuzzy logic does not explicitly calculate the selected class [10] but takes into account all parameters in each class, so that accuracy is expected to be more good when class determination. Example of calculating the value of the test data h1pada membership class 1st:

$$\mu_1(h1)$$

$$= \frac{\mu_{1,9}\|h1 - d_9\|^{-2} + \mu_{1,1}\|h1 - d_1\|^{-2} + \mu_{1,14}\|h1 - d_{14}\|^{-2}}{\|h1 - d_9\|^{-2} + \|h1 - d_1\|^{-2} + \|h1 - d_{14}\|^{-2}}$$

$$\mu_6(h1) = \frac{1\,.\,(0{,}5)^{-2} + 1\,.\,(1)^{-2} + 0\,.\,(1{,}73)^{-2}}{(0{,}5)^{-2} + (1)^{-2} + (1{,}73)^{-2}} = \frac{5}{5{,}33} = 0{,}938$$

5. Take the largest class membership value
6. Give the class with the largest membership value as a result of classification of test data

2.1.6 Information Extraction

After the classification results are obtained, the sentences in the test data table will be arranged into a new table, namely the extraction table in accordance with the class results of the classification. The following examples of extraction tables can be seen in Table 6.

Table 6. Table of Extraction Results

NO	NAMA INSTANSI	ALAMAT INSTANSI	NOMOR SURAT	TANGGAL
1	MUSYAWARAH GURU MATA PELAJARAN PENDIDIKAN AGAMA ISLAM (PAI) & BUDI PEKERTI SMP NEGERI KABUPATEN KEDIRI	Jalan Raya Turus No. 108 Desa Turus Kec. Gurah Kab. Kediri Email: mgmppaismpkabkediri@gmail.com Blog: mgmppaismpkabkediri.blogspot.com	Kediri, 4 Oktober 2016	Nomor : 005/ 397 /418.47.2.89.02/ 2016

2.2 System Testing

In testing this system, the data used were 400 training data with 75 test data obtained from the total preprocessing results of 60 incoming letter documents used. There are two tests carried out, namely:

2.2.1 The Confusion Matrix Test

number of training data used in this test is 400 data with the same distribution in each class totaling 50 data and test data as many as 75 data. Confusion matrix testing is done to determine the accuracy of the Fuzzy K-NN classification [11] and compare it with the K-NN accuracy and to find out the optimal k value. The results of the experiment testing the effect of K value on accuracy can be seen in Table 7.

Table 7. Test Results Confusion Matrix

No.	*k*	FKNN				KNN
		Precision	Recall	F-Score	Accuracy	Accuracy
1	1	88.75%	88.14%	86.50%	88.00%	88.00%
2	2	91.25%	90.28%	89.84%	90.67%	89.33%
3	3	90.00%	89.57%	88.10%	89.33%	85.33%
4	4	88.75%	88.10%	86.89%	88.00%	85.33%
5	5	90, 00%	89.57%	88.10%	89.33%	88.00%
6	6	90.00%	89.57%	88.10%	89.33%	88.00%
7	7	88.75%	88.01%	86.87%	88.00%	86.67%
8	8	88.75%	88.01%	86.87%	88.00%	85.33%
9	9	88.75%	88.01%	86.87%	88 00%	85.33%
10	10	88.75%	88.01%	86.87%	88.00%	84.00%
11	11	88.75%	88.01%	86.87%	88.00%	84.00%
12	12	87.50%	86.53%	85.66%	86.67%	84.00%
13	13	88.75%	88.01%	86.87%	88.00%	84.00%
14	14	88.75%	88.01%	86.87%	88.00%	84.00%
15	15	88.75%	87.39%	87.18%	88.00%	85.33%
16	16	88.75%	87.39%	87, 18%	88.00%	84.00%
17	17	88.75%	87.39%	87.18%	88.00%	85.33%
18	18	88.75%	87.39%	87.18%	88.00%	84.00%
19	19	86.25%	85.95%	83.77%	85.33%	84.00%
20	20	87.50%	87.09%	85.02%	86.67%	84.00%
21	21	86.25%	85.95%	83.77%	85.33%	84.00%
22	22	86.25%	85.95%	83.77%	85.33%	85.33%
23	23	86.25%	85, 95%	83.77%	85.33%	85.33%
24	24	86.25%	85.95%	83, 77%	85.33%	85.33%
25	25	86.25%	85.95%	83.77%	85.33%	85.33%
Average		88.30%	87.61%	86.31%	87.52 %	85.17%

of Table 7. obtained graphs of the calculation of precision, recall and f1-score which can be seen in Figure 9.

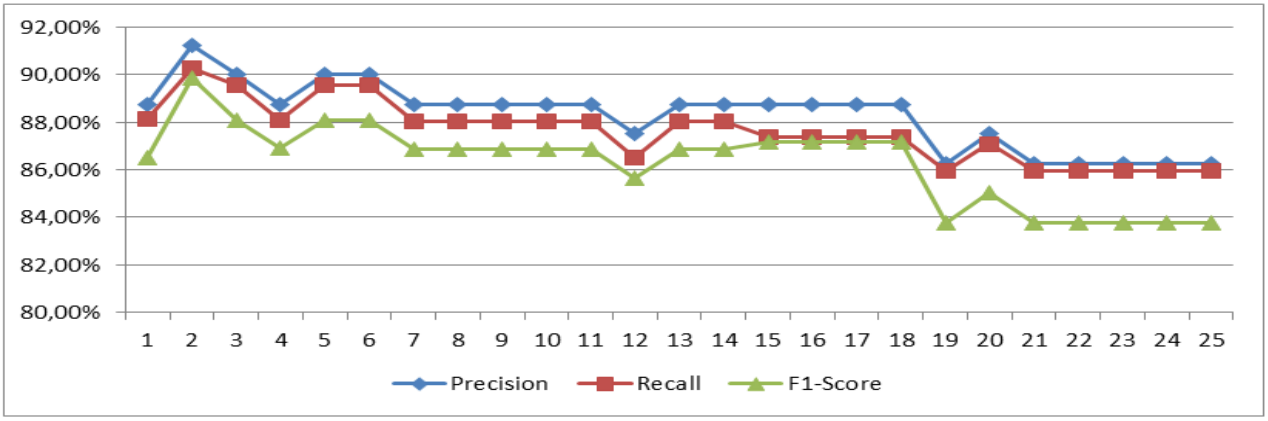
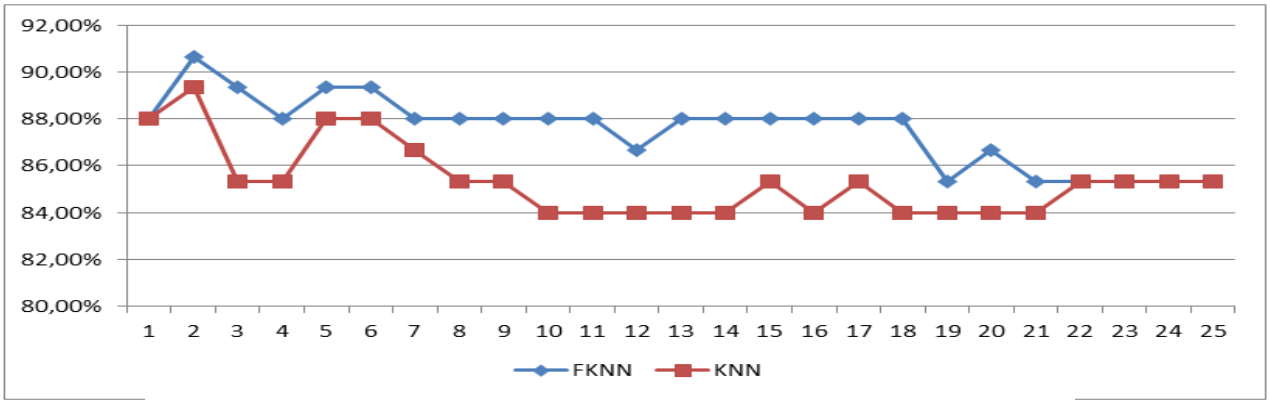


Figure 9. Graph of Calculation of Precision, Recall and F1-score

From Table 7. also obtained a graph of the comparison of the accuracy of Fuzzy K-NN and K-NN in Figure 10.



Gambar 10. Grafik Perbandingan FK-NN dan K-NN

Judging from the results in Table 7. and in the graph in Figure 9. and Figure 10. after the value of k = 21 until the value of k = 25 there has been no significant change so the use of k value stops at number 25. Then the average value of precision is obtained , recall, f1-score, and accuracy for the Fuzzy K-NN classification is worth 88.30%, 87.61%, 86.31%, and 87.52% with the highest precision, recall, f1-score and accuracy obtained at the time the value of k = 2 is 91.25%, 90.28%, 89.84%, and 90.67%. While the average accuracy of KNN is 85.17% with the highest accuracy value obtained when the value of k = 2 is 89.33%. So that the most optimal k value is 2.

2.2.2 Training Data Composition Testing

In testing the effect of training data composition on accuracy, the training data used were 240 training data with different compositions. This test aims to determine whether the composition of the training data is different, the accuracy produced in this study will have an effect or not. The 9 compositions that will be tested can be seen in Table 8.

Table 8 Composition of Training Data

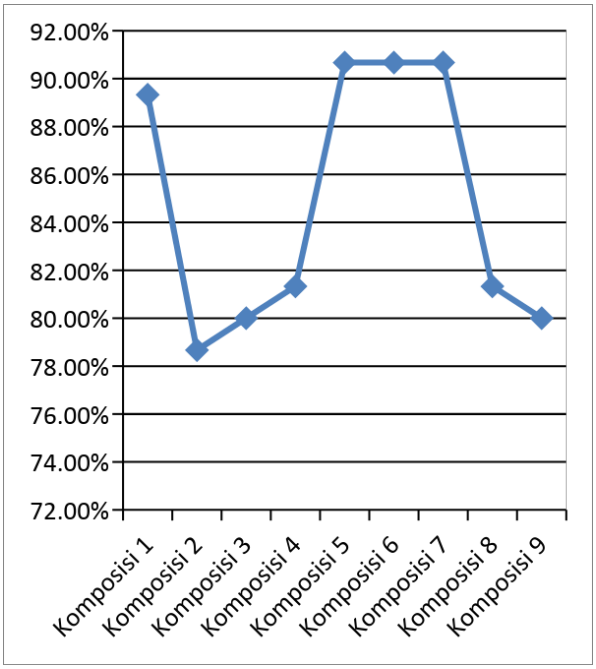
		Total Data Practice Class								TOT
		1	2	3	4	5	6	7	8	
Composition Data Training	1	30	30	30	30	30	30	30	30	240
	2	40	35	35	30	30	25	25	20	240
	3	20	40	35	35	30	30	25	25	240
	4	25	20	40	35	35	30	30	25	240
	5	25	25	20	40	35	35	30	30	240
	6	30	25	25	20	40	35	35	30	240
	7	30	30	25	25	20	40	35	35	240
	8	35	30	30	25	25	20	40	35	240
	9	35	35	30	30	25	25	20	40	240

The results of testing the composition of training data can be seen in Table 9. and the graph in Figure 11.

Table 9. Accuracy Calculation Results in Accordance with Training Data Composition

No.	Composition of Training Data	Accuracy
1	Composition 1	89.33%
2	Composition 2	78.67%
3	Composition 3	80.00%
4	Composition 4	81.33%
5	Composition 5	90.67%
6	Composition 6	90.67%
7	Composition 7	90,67%
8	Composition 8	81.33%
9	Composition 9	80.00%
Average		84.74%

Figure 11. Graph of Training Data Composition Testing



Based on Table 9. and Figure 11., the accuracy produced by the composition of training data that has classes the predominance is that classes 4.5 and 6 have higher accuracy than other compositions, while those with no grade 4.5 and 6 as the dominant class in their composition have decreased accuracy. This is due to the ability of the system to classify data with similar weight values (classes 4.5, and 6) experiencing a decrease due to reduced class training in the test data. Therefore, it is necessary to add features so that the weighting value can be more specific and avoid the same weighting values between classes, so as to minimize the decrease in accuracy due to reduced training data in certain classes.

3. CLOSING CONCLUSION

3.1 Conclusion

Conclusion from the results of the research is that the Fuzzy K-NN classification method can be applied in the extraction of document information and is quite capable of handling the characteristics of non-identical data such as incoming letters proven by the results of good accuracy and f1-score. Besides that the Fuzzy K-NN method has better accuracy compared to the K-NN classification method. Information extraction testing of incoming mail documents using the Fuzzy K-NN algorithm as a classification method produces the greatest accuracy of 90.67% at k = 2.

3.2 Suggestions

Within the existing limitations, some suggestions that can be given are using better conversion media. Because in this study there is still noise in the extraction results so that reading information can occur errors. This is due to an error in the conversion process in this study using the pdftables API. Then it is recommended to use more features so that the weighting value can be more specific. The goal is to minimize the occurrence of a decrease in the accuracy of the Fuzzy K-NN classification due to reduced training data that has nearly the same weight value.

REFERENCES

[1] SM Weiss, N. Indurkhya, F. Damerau, and T. Zhang, Text Mining Methods for Analyzing Unstructured Information, Springer, 2004.
[2] J. Han, M. Kamber, and J. Pei, Data Mining Concepts and Third Edition Techniques, Morgan Kaufmann Publishers, 2012.
[3] Minister of Administrative Reform and Bureaucratic Reform of the Republic of Indonesia, Guidelines for Managing Officials of Government Agencies, Ministry of Administrative Reform and Bureaucratic Reform of the Republic of Indonesia, 2012.
[4] Ismaya , Agny, "Rules-Based Information Extraction Algorithm", National Journal of Electrical Engineering and Information Technology, Vol. 03, No. 04, pp. 242–247, 2014.
[5] Hermawan, "Archive and Administration Data Collection Application in the Jeruk Village Using Java Netbeans", Muhammadiyah University Surakarta, 2013.

[6] J. Zhang, Y. Niu, and H. Nie, "Web Document Classification Based on Fuzzy k-NN Algorithm ", Int. Conf. Comput. Intell. Secur., Pp. 193-196, 2009.
[7] SD Nugraha, RRM Putri, and RC Wihandika, "Application of Fuzzy K-Nearest Neighbor (FK-NN) in Determining the Nutritional Status of Toddlers", Journal of the Development of Information Technology and Computer Science, vol. 1, no. 9, pp. 925–932, 2017.
[8] J. Han and M. Kamber, Data Mining Concepts and Techniques Second Edition, Morgan Kaufmann Publishers, 2006.
[9] JM Keller and MR Gray, "A Fuzzy K-Nearest Neighbor Algorithm", IEEE Trans. Syst. Man Cybern., Vol. SMC-15, no. 4, PP. 580-585, 1985.
[10] I., Nelly, "Model of Behavior of Walking Agents Using Fuzzy Logic", Journal of Computers and Information (KOMPUTA), vol. Volume. 1, no. Edition. I, pp. 37–43, 2012.
[11] M. Sokolova and G. Lapalme, "A Systematic Analysis of Performance Measures for Classification Tasks", Information Processing & Management., Vol. 45, no. 4, pp. 427-437, 2009.