HASIL CEK_Riadi_Cyberbullying; social media; instagram; Kmeans; clustering

by Imam Riadi Cyberbullying Analysis On Instagram_juita_10_2

Submission date: 13-Dec-2022 11:39AM (UTC+0700)
Submission ID: 1979873680
File name: Cyberbullying_Analysis_on_Instagram_Using_K-Means_Clustering.pdf (1.66M)
Word count: 6106
Character count: 32137

JUITA: Jurnal Informatika e-ISSN: 2579-8901; Vol. 10, No. 2, November 2022

Cyberbullying Analysis on Instagram Using K-Means Clustering

Ahmad Muhariya[1], Imam Riadi[2], Yudi Prayudi[3]
[1,3]Department of Informatics, Universitas Islam Indonesia Yogyakarta, Indonesia
[2] Department of Information System, Universitas Ahmad Dahlan Yogyakarta, Indonesia
[1]ahmad.muhariya@students.uii.ac.id [2]imam.riadi@is.uad.ac.id [3]prayudi@uii.ac.id

Abstract— Social Media, in addition to having a positive impact on society, also has a negative effect. Based on statistics, 95 percent of internet users in Indonesia use the internet to access social networks. Especially for young people, Instagram is more widely used than other social media such as Twitter and Facebook. In terms of cyberbullying cases, cases often occur through social media, Twitter, and Instagram. Several methods are commonly used to analyze cyberbullying cases, such as SVM (Support Vector Machine), NBC (Naïve Bayes Classifier), C45, and K-Nearest Neighbors. Application of a number of these methods is generally implemented on Twitter social media. Meanwhile, young users currently use Instagram more social media than Twitter. For this reason, the research focuses on analyzing cyberbullying on Instagram by applying the K-Mean Clustering algorithm. This algorithm is used to classify cyberbullying actions contained in comments. The dataset used in this study was taken from 2019 to 2021 with 650 records; there were 1827 words and already had labels. This study has successfully classified the tested data with a threshold value of 0.5. The results for grouping words containing bullying on Instagram resulted in the highest accuracy, which is 67.38%, a precision value of 76.70%, and a recall value of 67.48%. These results indicate that the k-means algorithm can make a grouping of comments into two clusters: bullying and non-bullying.

Keywords: Cyberbullying; social media; instagram; Kmeans; clustering;

I. INTRODUCTION

Social media is a form of technological progress that can easily affect society regarding perspective, lifestyle, and culture. Indonesia is a country with the 3rd most significant number of social media users in the world. As many as 120 million people in Indonesia use mobile devices such as smartphones or tablets to access social media, with a penetration of 45 percent. Online activity on social media via smartphones reaches 37 percent in a week [1]. According to data from the Ministry of Communication and Information, the number of internet users in Indonesia has reached 175.5 million people. Social media can have not only a positive impact on society but also a negative impact. One of them is using social media by certain parties which leads to cybercrime [2]. From this data, 95 percent of internet users actively access social networks such as Instagram, Facebook, and Twitter. In this case, the statistics also state that Instagram is more dominant among young people than Twitter. Therefore, Instagram is the most frequently used social media for posts that lead to cases of cyberbullying.

According to a survey from the Anti-bullying agency Ditch The Label, cyberbullying is negative comments on posts, personal messages that are not good, and making fun of others. Cyberbullying can be defined as behavior that harasses, insults, threatens, demeans, or harms someone continuously by utilizing technology the internet, and social media. Cyberbullying is more painful than physical violence [3]. Actions taken by someone who does cyberbully can be grouped into several different levels such as flaming, harassment, cyberstalking, slander, exclusion, trolling, impersonation and deception [4]. While at the 3rd International Conference 2018, Cyberbullying was categorized into three types namely threats, curses, and sexual [5].

Written words/text that contains hate or bullying are actually challenging to detect by the system. This is because of the many languages and accents used when pronouncing a text. For this reason, research conducted by some previous researchers uses a one-language approach, and data collection is done manually. However, in this study, a bilingual method is used, and the use of APIs for data collection purposes is used [6].

There are six stages that need to be done to conduct a cyberbullying investigation. The stages in question are building a corpus, pre-processing, identifying the features of each post that are checked with TF-IDF and a count vectorizer, modeling and labeling training, classification of cyberbullying, and corpus cyberbullying. However, the six stages have not solved the problem of digital evidence for the existence of cyberbullying. In addition to the general stages in cyberbullying investigations, there are other necessary

stages namely collecting data, pre-processing, cleaning the data, then classifying the data to compare sentiment data [7].

From the amount of literature obtained, many methods and algorithms are used for cyberbullying analysis including SVM (Support Vector Machine), NBC (Naïve Bayes Classifier), C45, and K-Nearest Neighbors. However, the algorithm is implemented chiefly on Twitter not Instagram. Similar research with the Instagram object has succeeded in grouping Instagram users based on the suitability of specific hashtags in Instagram text. In this case, the k-means and TF-IDF algorithms are used as the main features for the grouping [8]. However, applying the k-means clustering algorithm for cyberbullying analysis on Instagram has never been done [9]. The k-means method is used to classify cyberbullying actions in comments. K-means is a method that is partitional data. K-means works by dividing the data set into non-overlapping clusters so that each data is in exactly one group. Determination of the results of cluster values seen from the closest distance between the data and the centroid [10]. K-Means clustering aims to optimize a function to calculate the space distance between objects and the centroid (middle point) of the cluster [11]. Thus, it is hoped that the use of K-means clustering can also be done to group comments on Instagram that contain elements of bullying. The contribution of this research is in the application of the K-means clustering algorithm combined with the tf-idf weighting for cyberbullying analysis on Instagram. The implementation of this algorithm for cyberbullying analysis on Instagram has never been done.

II. METHOD

This study uses data from Instagram comments containing positive and negative elements. The stages in this research are preprocessing and term weighting, and then the clustering process is carried out, and finally, the evaluation process. The research steps can be seen in Fig. 1.

A. Preprocessing

This preprocessing stage aims to eliminate noise in the data to be processed which can affect clustering results. This stage is carried out so the grouping process can be done correctly. According to [12], the preprocessing set consists of several steps, namely:

- Case Folding is a step to change all sentences in lowercase [13].
- Tokenizing is a step to change the form of a sentence into several words in a sentence and remove punctuation marks in a sentence [14].
- Stopword is the process of removing essential words into a sentence obtained from the previous step. The words in the stopword are already listed in the stopword list dictionary [15].
- Normalization is a step to change from non-standard to standard words [16].
- Stemming is removing words into basic words [17].

In general, five preprocessing steps are often used but there is also a lemmatization step. However in this study the concept of stemming is used because the dataset used is in the form of the Indonesian language, and the library used is Sastrawi. Meanwhile, it is not easy to find a library in Indonesian using lemmatization.

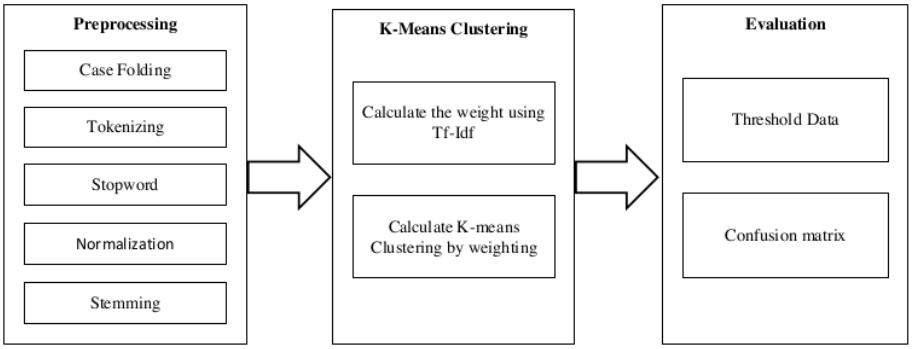

Fig. 1 Preprocessing stages

B. Term Weighting

The output of the preprocessing stage is the generation of a set of terms or words. Then the next step is to carry out the term weighting process, namely giving weight or value that indicates the importance of a term or word to the document. Calculation of the weight of each term or word is sought in each document with the aim of being able to find out the availability and similarity of a term or word in the document [18]. The more terms that appear in the document collection, the higher the value or weight of the term. After the weighting stage is complete, then proceed to the clustering process. In the term weighting stage, the method used in weighting is the Tf-Idf method.

Term Frequency (Tf) is the determination of a document's weight by calculating the frequency of occurrence of a term in the document. The more often a term or word appears, the higher the document weight for that term or word and vice versa.

Inverse Document Frequency (IDF) focuses on the occurrence of the term in the entire text collection in comments. In Idf, terms rarely appear in the entire term collection and are considered more valuable. Inverse Document Frequency (IDF) is calculated using (1).

$$\text{Idf} = \log\left(\frac{\textit{the total number of documents in the collection}}{\textit{number of documents containing the term}}\right) \quad (1)$$

Thus, the formula for calculating Tf-Idf combines the Tf formula with the Idf formula by multiplying the Term Frequency (Tf) value with the Inverse Document Frequency (Idf).

C. Clustering

Clustering is the stage to apply data mining techniques namely groups of the same object are combined to form a cluster and cluster this different from objects in other clusters [19]. At this stage, to group cyberbullying actions in comments that contain positive and negative elements, the following equation is used [20]:

- The vector object obtained from the weighting process is allocated, and then the centroid is determined randomly.
- The distance between the centroid and the object or term is calculated using Euclidian Distance. To calculate it by (2). Euclidean Distance formula:

$$d_{ij} = \sum_{k=1}^{n}\left(x_{ik} - x_{jk}\right)^2 \quad (2)$$

With:
= degree of difference
n = number of vectors
x_{ik} = input image vector
x_{jk} = vector of comparison/output image

- If the centroid changes again, the process returns to step 3 by determining the position of the new centroid using (3)

$$v = \frac{\sum_{i=1}^{n} x_i}{n};\ \text{i=1, 2, 3, …, n} \quad (3)$$

with:
v = centroid of the cluster
x_i = i-th object
n = number of objects/number of objects that are members of the cluster

If the position of the centroid does not change again, the clustering process is complete and the results obtained are grouping objects in certain categories based on the nearest centroid.

D. Evaluation

Evaluation is the stage to apply the confusion matrix to see how effective and accurate the performance of the model has been made. The confusion matrix will contain information about TP, FP, TN, and FN which will be useful for viewing clustering results. This study will divide the data used by 10-fold cross-validation to evaluate the performance of the model or algorithm. The higher the number generated in the confusion matrix, the better the resulting model and indicates a higher accuracy for the clustering process. Three points will be used as benchmarks for the value of the resulting model that is Accuracy, Precision, and Recall. Accuracy, Precision, and Recall can be assigned numerical values by using percentage calculations (1-100%) or by using numbers between 0 -1. The recommendation system will be considered good if the value of Accuracy, Precision, and Recall is high.

E. Case Simulation

A case simulation for the application of Cyberbullying Analysis on Instagram Social Media Services Using K-Means Clustering can be seen in Fig. 2. In this case simulation, the victim is assumed to report the case to the investigator and then follow up as a cyber investigation.

III. RESULTS AND DISCUSSION

The data used when using the crawling technique will lead to a less objective assessment of the manual labeling, thus using existing datasets from research sources that have been used. The cyberbullying dataset on Instagram social media used in this study has been used in previous studies using a different method,

Fig. 2 Case simulation

namely classification. There are several datasets sourced from:

- https://github.com/rizalespe/Dataset-Sentimen-Analisis-Bahasa-Indonesia/blob/master/dataset_komentar_instagram_cyberbullying.csv
- https://www.kaggle.com/datasets/cttrhnn/cyberbullying-bahasa-indonesia.

In this study, the dataset is sourced from www.kaggle.com because it is up to date, where this data was collected from 2019 to 2021 and has a label. The dataset in this study is open public, containing 650 records and 1827 words, and already has a label or category. There are 5 attributes and 1 class attribute contained in the dataset, as shown in Table I.

Of the five attributes in Table I, only the Instagram name, comments, and label or category attributes are used for this research. Table II is the metadata of the dataset used in this study.

TABLE I
CYBERBULLYING DATA ATTRIBUTES ON INSTAGRAM

No.	Attributes
1	Instagram Account
2	Comment
3	Category
4	Post Date
5	Victim's Account Name

TABLE II
COMMENTS ON INSTAGRAM

No	Instagram Account	Comments
1	@delliananda	*"Kaka tidur yaa, udah pagi, gaboleh capek2"*
2	@fenninbl	*"makan nasi padang aja begini badannya"*
3	@abdurahmanshq	*"yang aku suka dari dia adalah selalu cukur jembut sebelum manggung"*
4	@najla.yoo	*"Hai kak Isyana aku ngefans banget sama kak Isyana.aku paling suka lagu kak Isyana itu lagu tetap didalam jiwa"*
5	@dessy_____	*"Manusia apa bidadari sih herann deh cantik terus 😊♥"*
6	@e.fril	*"@ayu.kinantii isyan skrg berubah ya:(baju nya nakal"*
7	@bahasa.bayi.planet	*"Gemesnya isyan kayak tango, berlapis lapis ciaaaa"*
8	@khanayarudinita	*"Makin jelek aja anaknya, padahal ibu ayahnya cakep2"*
9	@reniaulia225	*"Kok anaknya kayak udah tua gitu ya mukanya kk tasya"*
10	@nurjanah.hani	*"Muka anak nya ko tua banget yaa.. GK ngegemesin GK ada lucu2nya"*
...	...	...
650	@dikha.wirasathya	*"Inimah bukan main alat musik lagi. Olahraga jari dan kaki ini mah"*

A. *Preprocessing*

In this research, 650 records containing 1827 words have been calculated. The collected comments are then preprocessed to speed up and simplify the grouping. Before the dataset is entered into the model to be proposed, there are several stages carried out at the data preprocessing stage, namely:

1) Case Folding. The results of the changes in the case folding stage can be seen in Table III. In the table, the first column contains the number, the second column contains the Instagram name, and the third column contains the comments expressed on Instagram. The difference between Table 2 and Table 3 is in changing the capital letter at the beginning of the word to lowercase. An example is the word "Kaka" becomes "kaka". The results of case folding are used to convert comments into small words, making it easier to compose text.

2) Tokenizing. The results at the tokenizing stage can be seen in Table IV. In that table, there is a column containing comments expressed on Instagram. The difference between Table III and Table IV is the change in the comma "," while in each sentence the sign is a word separator. The tokenizing results divide sentences

into words written by the alleged perpetrator in a comment on Instagram. This process will make it easier to remove non-essential words in the comment.

3) Stopword. The results of the changes made at the stopword stage can be seen in Table V. The difference between Table IV and Table V is in the presence of words written in italics and bold, such as the word *"aja", "yang", "aku", "dari", "dia", "adalah", "selalu", "sebelum", "sama", "paling", "itu", "tetap", "apa", "sih", "terus", "ya", "nya"*, and others. These words are removed from the comments sentence.

TABLE III
CASE FOLDING

No	Instagram Account	Comments
1	@dellianananda	*"kaka tidur yaa, udah pagi, gaboleh capek2"*
2	@fenninbl	*"makan nasi padang aja begini badannya"*
3	@abdurahmanshq	*"yang aku suka dari dia adalah selalu cukur jembut sebelum manggung"*
4	@najla.yoo	*"hai kak isyana aku ngefans banget sama kak isyana.aku paling suka lagu kak isyana itu lagu tetap didalam jiwa"*
5	@dessy_____	*"manusia apa bidadari sih herann deh cantik terus 😊♥"*
6	@e.fril	*"@ayu.kinantii isyan skrg berubah ya😕baju nya nakal"*
7	@bahasa.bayi.planet	*"gemesnya isyan kayak tango, berlapis lapis ciaaaa"*
8	@khanayarudinita	*"makin jelek aja anaknya, padahal ibu ayahnya cakep2"*
9	@reniaulia225	*"kok anaknya kayak udah tua gitu ya mukanya kk tasya"*
10	@nurjanah.hani	*"muka anak nya ko tua banget yaa.. gk ngegemesin gk ada lucu2nya"*
...	...	...
650	@dikha.wirasathya	*"inimah bukan main alat 265ahas lagi. olahraga jari dan kaki ini mah"*

4) Normalization. The results of the changes made at the normalization stage can be seen in Table VI. The difference between Table V and Table VI is in the change from non-standard words to standard words, such as the word *"yaa"* changing to *"iya"*, *"skrg"* changing to *"sekarang"*, *"kayak"* changing to *"seperti"*, *"gk"* changing to *"tidak"*, and others. The words will be changed from the comment.

TABLE IV
TOKENIZING

No	Instagram Account	Comments
1	@dellianananda	*'kaka', 'tidur', 'yaa', 'udah', 'pagi', 'gaboleh', 'capek'*
2	@fenninbl	*'makan', 'nasi', 'padang', 'aja', 'begini', 'badannya'*
3	@abdurahmanshq	*'yang', 'aku', 'suka', 'dari', 'dia', 'adalah', 'selalu', 'cukur', 'jembut', 'sebelum', 'manggung'*
4	@najla.yoo	*'hai', 'kak', 'isyana', 'aku', 'ngefans', 'banget', 'sama', 'kak', 'isyanaaku', 'paling', 'suka', 'lagu', 'kak', 'isyana', 'itu', 'lagu', 'tetap', 'didalam', 'jiwa'*
5	@dessy_____	*'manusia', 'apa', 'bidadari', 'sih', 'herann', 'deh', 'cantik', 'terus'*
6	@e.fril	*'kinantii', 'isyan', 'skrg', 'berubah', 'ya', 'baju', 'nya', 'nakal'*
7	@bahasa.bayi.planet	*'gemesnya', 'isyan', 'kayak', 'tango', 'berlapis', 'lapis', 'ciaaaa'*
8	@khanayarudinita	*'makin', 'jelek', 'aja', 'anaknya', 'padahal', 'ibu', 'ayahnya', 'cakep'*
9	@reniaulia225	*'kok', 'anaknya', 'kayak', 'udah', 'tua', 'gitu', 'ya', 'mukanya', 'kk', 'tasya'*
10	@nurjanah.hani	*'muka', 'anak', 'nya', 'ko', 'tua', 'banget', 'yaa', 'gk', 'ngegemesin', 'gk', 'ada', 'lucunya'*
...	...	...
650	@dikha.wirasathya	*'inimah', 'bukan', 'main', 'alat', 'musik', 'lagi', 'olahraga', 'jari', 'dan', 'kaki', 'ini', 'mah'*

TABLE V
STOPWORD

No	Instagram Account	Comments
1	@delliananda	*'kaka', 'tidur', 'yaa', 'udah', 'pagi', 'gaboleh', 'capek'*
2	@fenninbl	*'makan', 'nasi', 'padang', **'aja'**, 'begini', 'badannya'*
3	@abdurahmanshq	***'yang', 'aku'**, 'suka', **'dari', 'dia', 'adalah', 'selalu'**, 'cukur', 'jembut', **'sebelum'**, 'manggung'*
4	@najla.yoo	*'hai', 'kak', 'isyana', **'aku'**, 'ngefans', 'banget', **'sama'**, 'kak', 'isyanaaku', **'paling'**, 'suka', 'lagu', 'kak', 'isyana', **'itu'**, 'lagu', **'tetap'**, 'didalam', 'jiwa'*
5	@dessy_____	*'manusia', **'apa'**, 'bidadari', **'sih'**, 'herann', **'deh'**, 'cantik', 'terus'*
6	@e.fril	*'kinantii', 'isyan', 'skrg', 'berubah', **'ya'**, 'baju', **'nya'**, 'nakal'*
7	@bahasa.bayi.planet	*'gemesnya', 'isyan', 'kayak', 'tango', 'berlapis', 'lapis', 'ciaaaa'*
8	@khanayarudinita	***'makin'**, 'jelek', **'aja'**, 'anaknya', **'padahal', 'ibu'**, 'ayahnya', 'cakep'*
9	@reniaulia225	***'kok'**, 'anaknya', 'kayak', 'udah', 'tua', 'gitu', **'ya'**, 'mukanya', 'kk', 'tasya'*
10	@nurjanah.hani	*'muka', 'anak', 'nya', 'ko', 'tua', 'banget', 'yaa', 'gk', 'ngegemesin', 'gk', **'ada'**, 'lucunya'*
…	…	…
650	@dikha.wirasathya	*'inimah', **'bukan'**, 'main', 'alat', 'musik', **'lagi'**, 'olahraga', 'jari', 'dan', 'kaki', **'ini'**, 'mah'*

5) Stemming. The final stage is stemming, which uses sastrawi libraries, as shown in Table VII. The difference between Table VI and Table VII is in words written in italics and bold, such as the word "*badanya", "didalam", "berubah", "gemesnya", "anaknya", "ayahnya", "mukanya", "lucunya*", and others. The words will be changed from the comment.

After passing the preprocessing stage, the data output is the full document obtained is 1377 words.

TABLE VI
NORMALIZATION

No	Instagram Account	Comments
1	@delliananda	*'kaka', 'tidur', **'iya'**, 'sudah', 'pagi', 'gaboleh', 'capek'*
2	@fenninbl	*'makan', 'nasi', 'padang', 'badannya'*
3	@abdurahmanshq	*'suka', 'cukur', 'jembut', 'manggung'*
4	@najla.yoo	*'hai', 'kak', 'isyana', 'ngefans', 'banget', 'kak', 'isyanaaku', 'suka', 'lagu', 'kak', 'isyana', 'lagu', 'didalam', 'jiwa'*
5	@dessy_____	*'manusia', 'bidadari', 'herann', 'deh', 'cantik'*
6	@e.fril	*'kinantii', 'isyan', **'sekarang'**, 'berubah', 'baju', 'nakal'*
7	@bahasa.bayi.planet	*'gemesnya', 'isyan', **'seperti'**, 'tango', 'berlapis', 'lapis', 'ciaaaa'*
8	@khanayarudinita	*'jelek', 'anaknya', 'ayahnya', 'cakep'*
9	@reniaulia225	*'anaknya', **'seperti'**, 'sudah', 'tua', 'gitu', 'mukanya', 'kk', 'tasya'*
10	@nurjanah.hani	*'muka', 'anak', 'ko', 'tua', 'banget', **'iya'**, **'tidak'**, 'ngegemesin', **'tidak'**, 'lucunya'*
…	…	…
650	@dikha.wirasathya	*'inimah', 'main', 'alat', 'musik', 'olahraga', 'jari', 'kaki', 'mah'*

B. Term Weighting

After the preprocessing stage for 650 data, the next term is converted into vector data through the multiplication of Tf*Idf so that 1,377 terms or words are obtained. The following terms are obtained after the preprocessing results, as seen in Table VIII.

The next stage is the multiplication process between Term Frequency (Tf) and Inverse Document Frequency (Idf). The results of the Tf process can be seen in Table IX, while the Tf-Idf can be seen in Table X.

C. Clustering

After the preprocessing process is complete and converting the term into vector data through Tf*Idf multiplication, the clusterization process is carried out using K-Means. The results of system clustering using K-Means can be seen in Fig. 3 and Fig. 4.

TABLE VII
STEMMING

No	Instagram Account	Comments
1	@delliananda	*'kaka', 'tidur', 'iya', 'sudah', 'pagi', 'gaboleh', 'capek'*
2	@fenninbl	*'makan', 'nasi', 'padang', **'badan'***
3	@abdurahmanshq	*'suka', 'cukur', 'jembut', 'manggung'*
4	@najla.yoo	*'hai', 'kak', 'isyana', 'ngefans', 'banget', 'kak', 'isyanaaku', 'suka', 'lagu', 'kak', 'isyana', 'lagu', **'dalam'**, 'jiwa'*
5	@dessy_____	*'manusia', 'bidadari', 'herann', 'deh', 'cantik'*
6	@e.fril	*'kinantii', 'isyan', 'sekarang', **'berubah'**, 'baju', 'nakal'*
7	@bahasa.bayi.planet	***'gemes'**, 'isyan', 'seperti', 'tango', 'berlapis', 'lapis', 'ciaaaa'*
8	@khanayarudinita	*'jelek', **'anak'**, **'ayah'**, 'cakep'*
9	@reniaulia225	***'anak'**, 'seperti', 'sudah', 'tua', 'gitu', **'muka'**, 'kk', 'tasya'*
10	@nurjanah.hani	*'muka', 'anak', 'ko', 'tua', 'banget', 'iya', 'tidak', 'ngegemesin', 'tidak', **'lucu'***
…	…	….
650	@dikha.wirasathya	*'inimah', 'main', 'alat', 'musik', 'olahraga', 'jari', 'kaki', 'mah'*

Based on the information on the clustering results, as shown in Fig. 3, the results of the clustering are obtained. The result is that cluster 0 has 57 data records, and the data is in comments that contain non-cyberbullying elements.

Meanwhile, based on the information in Fig. 4, it is found that cluster 1 has 9 records of data, and the data is in comments that have elements of cyberbullying.

D. Evaluation

Testing the confusion matrix model to determine accuracy, precision, and recall is carried out to determine the level of accuracy of the clustering results obtained by the system. The data to be evaluated is the result of clustering with threshold values of 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0. The data is used to compare the results of clustering obtained by the system on different amounts of data. Before entering the dataset test, the data must first have a label. The following is a comparison of labeling on testing the 0.1 threshold value, which has been labeled with the cluster carried out by the system. If grouped according to the cluster, the results can be seen in Table XI.

TABLE VIII
TERM OF PREPROCESSING

No	Term	No	Term
1	*amin*	9	*adek*
2	*abai*	10	*adik*
3	*abal*	11	*aduh*
4	*activity*	12	*aduhh*
5	*adab*	13	*aduuh*
6	*adam*	14	aesthetic
7	*adaptasi*	…	…
8	*adek*	1377	*yutuber*

TABLE IX
RESULTS TERM FREQUENCY

Doc	*abal*	activity	*adab*	*amin*	*anak*	…	*yutuber*
1	0	0	0	0	0	…	0
2	0	0	0	0	0	…	0
3	0	0	0	0	0	…	0
4	0	0	0	0	0.442	…	0
5	0	0	0	0	0.500	…	0
6	0	0	0	0	0	…	0
7	0	0	0	0	0.378	…	0
8	0	0	0	0	0.378	…	0
9	0	0	0	0	0.258	…	0
10	0	0	0	0	0	…	0
…	…	…	…	…	…	…	0
650	0	0	0	0	0	…	0

TABLE X
RESULTS TF-IDF

Doc	*abal*	activity	*adab*	*amin*	*anak*	…	*yutuber*
1	0	0	0	0	0	…	0
2	0	0	0	0	0	…	0
3	0	0	0	0	0	…	0
4	0	0	0	0	0.370	…	0
5	0	0	0	0	0.305	…	0
6	0	0	0	0	0	…	0
7	0	0	0	0	0.232	…	0
8	0	0	0	0	0.221	…	0
9	0	0	0	0	0.154	…	0
10	0	0	0	0	0	…	0
…	…	…	…	…	…	…	0
650	0	0	0	0	0	…	0

id	cluster	text
1	cluster_0	makan nasi padang badan
2	cluster_0	anak tolol aktif bund
3	cluster_0	ancur hidup sumpah
4	cluster_0	cocok usaha kuliner fixs semoha lancat terusssss
5	cluster_0	muka sensual norak
6	cluster_0	monyet
7	cluster_0	suka heran warga suka ngehujat orang pahala dosa
8	cluster_0	cantik banget pakai baju gitu
9	cluster_0	cowok wajah culun cupu
10	cluster_0	putus tumbal apartment
11	cluster_0	palsu badane edit asli lemak
12	cluster_0	kontrak skincare buluk
13	cluster_0	selamat ulang sadar pansos amin
14	cluster_0	mbak cantik umur sehat
15	cluster_0	sumpah lengan banget

Fig. 3 Cluster 0 result

id	cluster	text
16	cluster_1	gaga kayak anjing
21	cluster_1	manusia anjing cari uang gitu kerja nyebokin sebentar gaji
28	cluster_1	wiii anjing lihat babi megang anjing
29	cluster_1	sikap najis ludah anjing
31	cluster_1	kasi anjing peluk sampahh
32	cluster_1	mukak tapir kuda anjing ganteng banget
37	cluster_1	kayak anjing kayak babi
38	cluster_1	kayak lonte
57	cluster_1	kultur kerja kayak bagus kapitalis banget petinggi kerja ngevlog endors moga g...

Fig. 4 Cluster 1 result

TABLE XI
GROUPING OF CLUSTER RESULTS WITH A THRESHOLD VALUE OF 0.1

Cluster	Many Comments	Category
Cluster 0	57 Comments	Non-Cyberbullying
Cluster 1	9 Comments	

In this case, the category of each cluster is determined by the researcher because the system does not know the category of each cluster. The system only generates clusters of each specified comment. To make it easier to calculate the values of accuracy, precision, and recall, relevant and irrelevant data are traced to the data that has been clustered, as can be seen in Table XII.

TABLE XII
CLUSTER SEARCH RESULTS WITH A THRESHOLD VALUE OF 0.1

		Actual (Real)		Total Prediction
		C1	C2	
Prediction (system)	C1	32	25	**57**
	C2	1	8	**9**
Actual Total		33	33	N=66

Based on Table XII that has been done, the accuracy values obtained are as follows: accuracy = $\frac{40}{66} \times 100\% = 60.61\%$. The precision and recall values of each cluster can be seen in Table XIII.

From Table XIII, it can be seen that cluster 1 has a fairly high level of precision. This means that in one cluster, the correct data obtained are more than the incorrect data. Meanwhile, cluster 0 has a lower level of precision than cluster 1. The following is a comparison of the values of accuracy, precision, and recall with a threshold of 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0, 8, 0.9, 1.0, as seen in Table XIV.

Table XIV and Fig. 5 show that in the test with a threshold value of 0.5, the highest accuracy, average precision, and recall values were obtained, respectively, 67.38% for accuracy and 76.70% for precision. As for the recall, it is 67.48% from the other data tests.

The information obtained from the clustering results is shown in Fig. 3 and Fig. 4. There are 3 columns presented in the two images. The first column displays the id, the second column displays the cluster, and the third column displays the Instagram text or comments. For example, cluster 1 has 9 data records. The data is in comments that have elements of cyberbullying. The research results that have been done are comments with id 16 indicated as cyberbullying comments. These results can be used as digital evidence for trial purposes.

Based on the results submitted using as many as 650 records, it can be seen that the K-Means algorithm can run as expected. The K-Means algorithm is one of the recommended algorithms and is widely used for analysis on Instagram social media [8]. This study uses the value of K = 2 from the number of clusters that have been determined. Data processing is done using rapidminer software. The study's results obtained information that the test with a threshold value of 0.5 resulted in cluster 0, consisting of 67 records. Meanwhile, the cluster 1 formed consists of 258 records. Other information obtained is related to the value of accuracy, average precision, and recall. The highest level is 67.38% for accuracy, 76,70% for precision, and 67.48% for recall. These results were obtained from testing other data.

TABLE XIII
PRECISION AND RECALL RESULTS WITH A THRESHOLD VALUE OF 0.1

Cluster	Category	Precision	Recall
Cluster 0	Non-Cyberbullying	56.14%	96.97%
Cluster 1	Cyberbullying	88.89%	24.24%
Average		**72.5%**	**60.61%**

TABLE XIV
COMPARISON OF ACCURACY, PRECISION AND RECALL VALUES

No	Threshold	Accuracy	Precision	Recall
1	0.1	60.61%	72.5%	60.61%
2	0.2	61.54%	78.26%	61.54%
3	0.3	61.22%	76.12%	61.23%
4	0.4	60.77%	76.38%	60.77%
5	0.5	67.38%	76.70%	67.48%
6	0.6	60.51%	75.75%	60.51%
7	0.7	59.91%	75.74%	59.91%
8	0.8	59.23%	75,64%	59,23%
9	0.9	59.04%	75,75%	59,00%
10	1.0	58.46%	75,63%	58,46%

Whereas in previous research [21], the results of cyberbullying testing on Instagram comments using the Support Vector Machine Classification Method with 400 data records produced the best accuracy rate of 90%, a precision of 94.44%, and a recall of 85%. So it can be said that research using the k-means clustering method produces a lower level of accuracy when compared to previous research using the Support Vector Machine method. Accuracy with the k-means method is lower than the SVM (Support Vector Machine) algorithm due to several things, namely the lack of combinations of data obtained, the number of datasets used, and the various data characteristics used.

There are obstacles faced in this study when analyzing text in Indonesian. This is because the complete Indonesian language corpus is still difficult to obtain. In addition, often comments on Instagram use a lot of regional languages and slang. This constraint causes the results obtained from the data mining process that has not yet reached the desired level of accuracy.

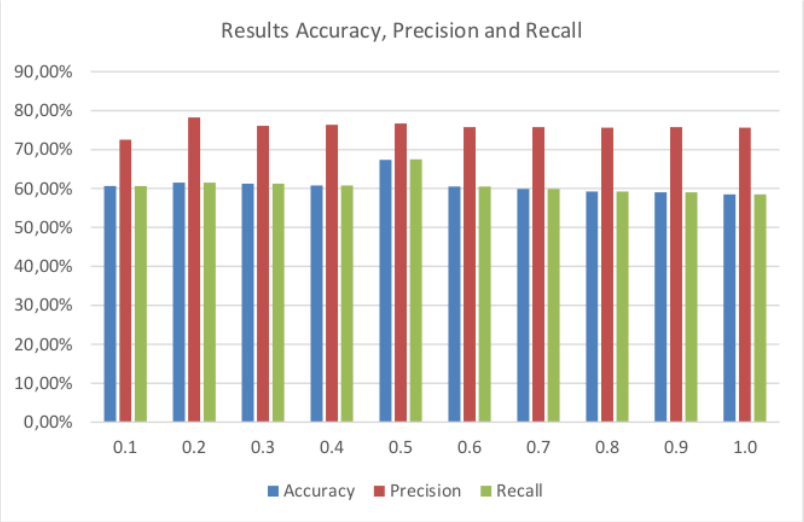

Fig. 5 Results of accuracy, precision and recall

IV. CONCLUSION

The conclusions and suggestions that can be drawn from this study are that several groups of data tested, with a threshold value of 0.5, had the highest accuracy, average precision, and recall compared to others, namely 67.38% for accuracy and 76,70% for precision. Meanwhile, the recall value is 67.48% from the other data tests. However, the resulting clusters have different accuracy, precision, and recall levels depending on the specified threshold. Thus, it can be said that the K-Means Algorithm can group comments into two clusters. Whereas in previous studies, cyberbullying testing on Instagram comments using the Support Vector Machine Classification Method resulted in the best quality accuracy of 90%, precision of 94.44%, and recall of 85%. So it can be said that the research using the k-means clustering method in this study resulted in a lower quality of accuracy when compared to previous studies using the Support Vector Machine method.

It is hoped that further research will add a data dictionary to normalize the words in the comments. This is recommended because there are still many words that are not standardized or contain regional languages. This will affect the results on the weighting of each word. In addition, further research can be carried out using the method to determine the best threshold value. It is also recommended that there is no need to do many experiments to determine how many of the best attributes will be used to produce the highest accuracy.

ACKNOWLEDGEMENT

This paper and the research that has been carried out is part of the Postgraduate Research Program Grant - Master Thesis Research organized by the Ministry of Education and Culture of the Republic of Indonesia with Research Assignment Number: 017/ST-DirDPPM/70/DPPM/PPS-PTM-KEMDIKBUDRISTEK/V/2022. The author expresses his gratitude to the Directorate of Research and Community Service (DPPM) and the Masters Program in Informatics, Universitas Islam Indonesia for the administrative support and research facilities that have been provided.

REFERENCES

[1] Ministry of Communications and Information Technolog, "Networking is Increasing, Indonesia Needs to Increase Cultural Values on the Internet" 2021. Https://Aptika.Kominfo.Go.Id/2021/09/Warganet-Meningkat-Indonesia-Perlu-Tingkatkan-Nilai-Budaya-Di-Internet/ (Accessed May 02, 2022) (In Indonesia Languange).

[2] L. Fazry And N. Cipta Apsari, "The Effect of Social Media on Cyberbullying Behavior Among Adolescents," *J. Pengabdi. Dan Penelit. Kpd. Masy.*, Vol. 2, No. 1, Pp. 28–36, 2021, [Online]. Available: Https://Ejournal.Bsi.Ac.Id/Ejurnal/Index.Php/Cakrawala/Article/Viewfile/3680/2624 (In Indonesia Languange).

[3] F. Tapia, C. Aguinaga, And R. Luje, "Detection Of Behavior Patterns Through Social Networks Like

Twitter, Using Data Mining Techniques As A Method To Detect Cyberbullying," *Appl. Softw. Eng. - Proc. 7th Int. Conf. Softw. Process Improv. Cimps 2018*, Pp. 111–118, 2019, Doi: 10.1109/Cimps.2018.8625625.

[4] P. W. Imam Riadi, Sunardi, "Investigating Cyberbullying On Whatsapp Using Digital Forensics," *Rekayasa Sist. Dan Teknol. Inf.*, Vol. 1, No. 10, Pp. 730–735, 2021 (In Indonesia Languange).

[5] O. C. Hang And H. M. Dahlan, "Cyberbullying Lexicon For Social Media," *Int. Conf. Res. Innov. Inf. Syst. Icriis*, Vol. December-2, 2019, Doi: 10.1109/Icriis48246.2019.9073679.

[6] R. Pawar And R. R. Raje, "Multilingual Cyberbullying Detection System," *Ieee Int. Conf. Electro Inf. Technol.*, Vol. 2019-May, Pp. 040–044, 2019, Doi: 10.1109/Eit.2019.8833846.

[7] F. Tapia And C. Aguinaga, "Detección De Patrones De Comportamiento A Través De Redes Sociales Como Twitter , Utilizando Técnicas De Minería De Datos Como Método Para Detectar El Acoso Cibernético Detection Of Behavior Patterns Through Social Networks Like Twitter , Using Data Minin," *2018 7th Int. Conf. Softw. Process Improv.*, Pp. 111–118, 2018, Doi: 10.1109/Cimps.2018.8625625.

[8] M. Habibi And P. W. Cahyo, "Clustering User Characteristics Based On The Influence Of Hashtags On The Instagram Platform," *Ijccs (Indonesian J. Comput. Cybern. Syst.*, Vol. 13, No. 4, P. 399, 2019, Doi: 10.22146/Ijccs.50574.

[9] M. Z. Naf'an, A. A. Bimantara, A. Larasati, E. M. Risondang, And N. A. S. Nugraha, "Sentiment Analysis Of Cyberbullying On Instagram User Comments," *J. Data Sci. Its Appl.*, Vol. 2, No. 1, Pp. 88–98, 2019, Doi: 10.21108/Jdsa.2019.2.20.

[10] J. Suryanto, "Comparative Analysis of 15 Daily Rainfall Groupings in DIY Province Using Fuzzy Clustering and K-Means Clustering," Vol. Xvi, Pp. 229–242, 2017 (In Indonesia Languange).

[11] M. H. Adiya And Y. Desnelita, "Application of K-Means Algorithm for Drug Data Clustering at Pekanbaru Hospital," Vol. 01, Pp. 17–24, 2019 (In Indonesia Languange).

[12] Rsa, "2016: Current State Of Cybercrime," P. 7, 2016, [Online]. Available: Https://Www.Rsa.Com/Content/Dam/Rsa/Pdf/2016/05/2016-Current-State-Of-Cybercrime.Pdf.

[13] H. S. Christopher D. Manning, Prabhakar Raghavan, *Introduction To Modern Information Retrieval*. 2009.

[14] I. Ruthven And M. Lalmas, "A Survey On The Use Of Relevance Feedback For Information Access Systems," *Knowl. Eng. Rev.*, Vol. 18, No. 2, Pp. 95–145, 2003, Doi: 10.1017/S0269888903000638.

[15] S. K. Sahu, S. Sarangi, And S. K. Jena, "A Detail Analysis On Intrusion Detection Datasets," *Souvenir 2014 Ieee Int. Adv. Comput. Conf. Iacc 2014*, Pp. 1348–1353, 2014, Doi: 10.1109/Iadcc.2014.6779523.

[16] R. Riyadulloh And A. Romadhony, "Normalization of Indonesian Text Based on Slang Dictionary Case Study: Tweet Product Gadgets on Twitter," *Eproceedings Eng.*, Vol. 8, No. 4, Pp. 4216–4228, 2021, [Online]. Available: Https://Openlibrarypublications.Telkomuniversity.Ac.Id/Index.Php/Engineering/Article/View/15246/14969 (In Indonesia Languange).

[17] E. W. Selberg, "Information Retrieval Advances Using Relevance Feedback," *Uw Dept Cse Gen. Exam*, 1997, [Online]. Available: Http://Www.Cs.Rpi.Edu/~Chapaa/Userskill/Paper/Generals.Pdf.

[18] P. Yugianus, H. S. Dachlan, And R. N. Hasanah, "Development of a Library Catalog Search System Using the Rocchio Relevance Feedback Method," *J. Eeccis*, Vol. 7, No. 1, Pp. 47–52, 2013 (In Indonesia Languange).

[19] I. Riadi, J. E. Istiyanto, A. Ashari, And Subanar, "Internet Forensics Framework Based-On Clustering," Vol. 4, No. 12, Pp. 115–123, 2013.

[20] M. Zulfadhilah, Y. Prayudi, And I. Riadi, "Cyber Profiling Using Log Analysis And K-Means Clustering," *Int. J. Adv. Comput. Sci. Appl.*, Vol. 7, No. 7, Pp. 430–435, 2016, Doi: 10.14569/Ijacsa.2016.070759.

[21] W. A. Luqyana, I. Cholissodin, And R. S. Perdana, "Analysis of Cyberbullying Sentiment on Instagram Comments With Support Vector Machine Classification Method," *J. Pengemb. Teknol. Inf. Dan Ilmu Komput. Univ. Brawijaya*, Vol. 2, No. 11, Pp. 4704–4713, 2018 (In Indonesia Languange).

HASIL CEK_Riadi_Cyberbullying; social media; instagram; Kmeans; clustering

ORIGINALITY REPORT

9%	4%	7%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Affi Nizar Suksmawati, Retantyo Wardoyo. "Prediction of O3 Concentration Level Using Fuzzy Non-Stationary Method", JUITA : Jurnal Informatika, 2022 Publication	3%
2	www.researchgate.net Internet Source	1%
3	Majid Rahardi, Afrig Aminuddin, Ferian Fauzi Abdulloh, Rizky Adhi Nugroho. "Sentiment Analysis of Covid-19 Vaccination using Support Vector Machine in Indonesia", International Journal of Advanced Computer Science and Applications, 2022 Publication	1%
4	commdis.telkomuniversity.ac.id Internet Source	1%
5	Desi Puspita, Sasmita Sasmita. "APPLICATION OF K-MEANS ALGORITHM IN GROUPING OF CITY TOURISM CITY PAGAR ALAM", Sinkron, 2022	1%

Publication

6	R Suwanda, Z Syahputra, E M Zamzami. "Analysis of Euclidean Distance and Manhattan Distance in the K-Means Algorithm for Variations Number of Centroid K", Journal of Physics: Conference Series, 2020 Publication	1 %
7	Fauzillah Indriani, Wahyu Wibowo. "Sentiment analysis data by ShopeePayLater Twitter's opinion using naïve Bayes classifier", AIP Publishing, 2022 Publication	1 %
8	Zulfadhilah, Muhammad, Yudi Prayudi, and Imam Riadi. "Cyber Profiling Using Log Analysis And K-Means Clustering", International Journal of Advanced Computer Science and Applications, 2016. Publication	1 %

Exclude quotes	On	Exclude matches	< 1%
Exclude bibliography	On		