**Accredited Ranking SINTA 2**
Decree of the Director General of Higher Education, Research and Technology, No. 158/E/KPT/2021
Validity period from Volume 5 Number 2 of 2021 to Volume 10 Number 1 of 2026

Published online on: **http://jurnal.iaii.or.id**

**JURNAL RESTI**
**(Rekayasa Sistem dan Teknologi Informasi)**
Vol. 7 No. 1 (2023) 185 - 192 ISSN Media Electronic: 2580-0760

# Covid-19 Fake News Detection on Twitter Based on Author Credibility Using Information Gain and KNN Methods

Nanda Ihwani Saputri[1], Yuliant Sibaroni[2], Sri Suryani Prasetiyowati[3]
[1,2,3]Informatika, Fakultas Informatika, Universitas Telkom
[1]nandaihwani@student.telkomuniversity.ac.id, [2]yuliant@telkomuniversity.ac.id, [3]srisuryani@telkomuniversity.ac.id

***Abstract***

*Twitter is one of the social media that is used as a tool to share various kinds of information about various kinds of things that are of concern to social media users. One of the information shared is information about COVID-19, which is known that the COVID-19 pandemic is currently spreading throughout the world at a very alarming rate. COVID-19 is an infectious disease caused by SARS-COV-2. The World Health Organization (WHO) claims that the spread of COVID-19 is supported by the spread of false/fake news. So to find out the truth of the news, a COVID-19 fake news detector is needed so that users don't fall for the hoaxes circulating. This study aims to classify COVID-19 news on Twitter based on author credibility. Credibility in question is a person's perception of the validity of information and is a multidimensional concept that is used as a means of receiving information to assess the source of communication. The method used in this research is Information Gain and KNN. KNN (K-Nearest Neighbor) is a supervised learning algorithm that works by classifying a set of data based on classified training data. Information Gain is used to ranking the most influential attributes, and KNN is used to classify data based on learning data taken from the nearest neighbors. The research consists of 6 main stages, namely data collection (crawling data), data preprocessing, feature extraction, feature selection, data split into training data and testing data, KNN stage, and data evaluation stage. The research carried out succeeded in obtaining an accuracy value of 91%, a correlation value between credibility and hoax of 0.115, and a p-value <0.005.*

*Keywords: twitter; fake news, COVID-19; credibility; KNN; information gain*

## 1. Introduction

The development of communication technology helps humans in sending and receiving information. Virtual communities are starting to form and shift traditional communities. The virtual community that we can meet is social media. Distribution of information in the form of online conversations can be done through social media. The most obvious participation and use of social media can be seen in social media such as Facebook and Twitter [1]. Based on a survey conducted by the Association of Indonesian Internet Service Providers, shows that internet penetration in Indonesia in 2020 will reach 171.17 million people, equivalent to 64.8% of Indonesia's total population [2]. Meanwhile, when a similar survey was conducted in 2017, 143 million internet users were found, equivalent to 54.7% of Indonesia's total population. This shows that there was an increase in the number of internet users from the previous year of 10.1%. This increase in internet access is in line with access to social media, both Facebook and Twitter.

Since its launch in 2006, Twitter has become one of the most popular social media platforms for sharing information, both personal information and a means of interaction in various parts of the world [3]. Dissemination of information on Twitter is done through making tweets. One of the information being disseminated is information about COVID-19, which is known that the COVID-19 pandemic is currently spreading throughout the world at a very alarming rate [4]. The World Health Organization (WHO) even claims that the spread of COVID-19 is supported by the spread of false/fake news [5].

Fake news about COVID-19 seems to spread very quickly on social media [6]. Similar trends have been seen during other epidemics, such as the Ebola, yellow fever, and Zika outbreaks [4]. This is a very worrying development because even a little exposure to fake news can cause public anxiety and distrust [5]. In addition, it is necessary to identify the creator or subject of fake news which will help eradicate a large number of fake news from its origin [7]. Generally, for news spreaders, in addition to the tweets that are made,

Accepted: 17-01-2023 | Received in revised: 03-02-2023 | Published: 06-02-2023

there is social media profile information such as username, location, user description, account creation time, and followers. Which information can be used to obtain fundamental complementary information for background checks as well as a basis for determining the credibility of news spreaders [7]. The credibility in question is a person's perception of the validity of information and is a multidimensional concept that is used as a means of receiving information to assess the source of communication [8].

Previous research presented case studies on credibility based on articles published by “Donald Trump”, “Mike Pence”, “Barack Obama”, and “Hillary Clinton” [7]. Next, it explains 4 case studies related to the credibility of the authors based on published articles. Then, dividing the case into 2 groups, namely republican and democratic. Most of the dataset from “Donald Trump” is evaluated incorrectly by 69%. For the "Mike Pence" dataset it is evaluated at 52%: 48% for true and false news. Meanwhile, the dataset "Barack Obama" and "Hillary Clinton" were considered correct, namely 76% and 73% [7].

Therefore, this study aims to classify news on Twitter tweets, using the Information Gain and KNN methods. This classification is intended to analyze whether the tweet is fake news (hoax) or non-hoax based on the author’s credibility.

## 2. Research Methods

This research was conducted using the Information Gain and KNN methods. The KNN method is used to classify data based on learning data taken from the nearest neighbors [9]. Meanwhile, Information Gain is used to ranking the most influential attributes [10]. This is done because feature selection (Information Gain) is an important part that can optimize classifier performance [10], [11]. It is known that the use of the Information Gain algorithm can reduce the vector dimensions in the dataset [11], [12].

### 2.1 Research data

The dataset used in this study was obtained from Kaggle, which is tweet data from Twitter, and was collected by Gabriel Preda using the Twitter API combined with Python scripts. Queries are run daily for a set period, to collect a larger sample of tweets. The tweets obtained contain the hashtag #covid19. Tweet data collection started on 25/7/2020 to 29/8/2020, with an initial batch of 17,000 data.

### 2.2 Research Framework

This study consists of 6 stages, namely data collection (data crawling), data preprocessing, feature extraction, feature selection (Information Gain), data split into training data and testing data, KNN stage, and data evaluation stage (Figure 1).

Table 1. Twitter Data with the Hashtag #COVID19

| User name | User followers | text |
|---|---|---|
| Time4fisticuffs | 9275 | @diane3443 @wdunlap @realDonaldTrump Trump never once claimed #COVID19 was a hoax. We all claim that this effort toâ€¦ https://t.co/Jkk8vHWHb3 |
| DIPR-J&K | 101009 | 25 July: Media Bulletin on Novel #CoronaVirusUpdates #COVID19 |
| hr bartender | 79956 | How #COVID19 Will Change Work in General (and recruiting, specifically) via/ @ProactiveTalent #Recruitingâ€¦ https://t.co/bjZxzGPMbK |
| Greater Visakhapatnam Municipal Corporation (GVMC) | 14357 | GVMC sanitation staff carrying out the regular sanitation activities to keep the city clean and prevent the spreadâ€¦ https://t.co/bkrQ5x6BCK |
| Fergus McPop | 1029 | Coronavirus Testing Fiasco: St MirrenÂ have pledged to undertake an "urgent review" of their Covid-19 testing procedâ€¦ https://t.co/bfel6gyXIq |
| Albert Trigg | 3956 | A review of the recent study (now retracted) which connected #5G with #COVID19 |
| .. | .. | .. |
| rugby365.com | 35049 | #ICYMI: New @wallabies coach Dave Rennie will have a few tough decisions to make in the build-up to this year's revâ€¦ https://t.co/trk0GDrii5 |

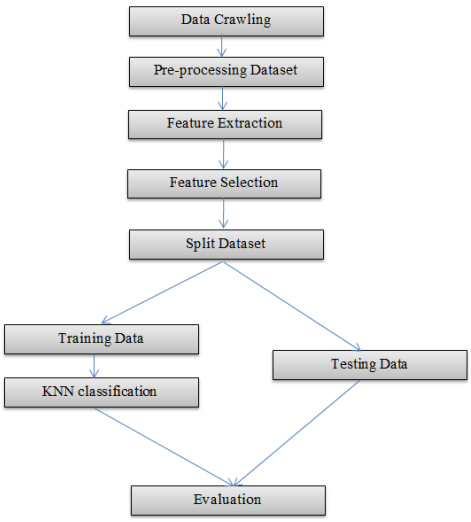


Figure 1. System Flowchart

In the data crawling process, tweets containing the hashtag #COVID19 are collected. Data is obtained from Twitter using the Twitter API combined with

DOI: https://doi.org/10.29207/resti.v7i1.4871

Python scripts. The data preprocessing process is carried out to process the dataset to facilitate classification, consisting of case folding, remove punctuation, tokenization, lemmatization, and stop word removal stages. Furthermore, feature extraction is carried out using TF-IDF calculations followed by feature selection using Information Gain on data that has gone through the preprocessing stage. Split datasets are used to develop statistical models and evaluate the performance of machine learning models. In this case, the dataset is divided into 2 parts, namely training data and testing data, with a scenario of 70%:30% and 60%:40% data sharing. After that, the training data and testing data are calculated using the Euclidean distance equation to find the shortest distance. The next process is measuring system performance (evaluation) which is carried out using a confusion matrix, such as calculating precision, recall, f1-score, and accuracy values.

### 2.3 Preprocessing Dataset

Dataset preprocessing is a process used in processing data that is not under the system and can interfere with results when processing data. In the classification of news that uses the data type in the form of text, there are several kinds of processes carried out, including case folding, remove punctuation, tokenization, lemmatization, and stop word removal. Case folding is used to standardize all characters in the text from uppercase to lowercase, remove punctuation is used to remove characters other than letters and punctuation. Tokenization is useful for separating each text in the form of sentences or paragraphs into words (tokens). Lemmatization is used to change words with the same meaning into one form. Stopword removal is useful for removing words that appear frequently but have no information in the text [13]. The results of the preprocessing process can be seen in Table 2.

Table 2. Preprocessing Data

| Voices from the Belt and Road: COVID rap song alerts to needed precautions when returning to work | |
|---|---|
| Preprocessing | text |
| *case folding* | voices from the belt and road: covid rap song alerts to needed precautions when returning to work |
| *remove punctuation* | voices from the belt and road covid rap song alerts to needed precautions when returning to work |
| *stop word removal* | Voices belt road covid rap song ales needed precautions returning work |
| *lemmatization* | voice belt road covid song ale need precaution return work |
| *tokenization* | 'voice' 'belt' 'road' 'covid' 'song' 'ale' 'need' 'precaution' 'return' 'work' |

### 2.4 Feature Extraction

Feature extraction is the process of turning raw data into processable numeric features while preserving the information in the original data set. Feature extraction produces better results than applying machine learning directly to raw data. The feature extraction used is TF-IDF. TF-IDF is a method used to evaluate the importance of a word in a document [14]. Term frequency (TF) is calculated as the number of times the term occurs in the document with the total number of words in the document. IDF (Inverse Document Frequency) is used to calculate the importance of a term. Calculations from TF-IDF can be seen in formulas (1), (2), and (3). From this process, as many as 5400 features were obtained which can be seen in Table 3.

$$tf(t,d) = \frac{f_{t,d}}{\sum_{t' \in d} f_{t',d}} \quad (1)$$

$$idf(t,D) = \log \frac{N}{|\{d \in D : t \in D\}|} \quad (2)$$

$$tfidf(t,d,D) = tf(t,d) \cdot idf(t,D) \quad (3)$$

Where $f_{t,d}$ is the raw number of terms in the document, i.e. the number of times term t appears in document d. N is the number of documents in the corpus. |{d∈D: t∈D}| is the number of documents in which the term t appears. If the term is not in the corpus, it results in a division by zero.

Table 3. Feature Extraction

| Words (5400 features) | | |
|---|---|---|
| 'covid' | 'bullshit' | 'additional' |
| 'case' | 'bundle' | 'amended' |
| 'still' | 'calculation' | 'amercians' |
| 'response' | 'cardiovascular' | 'believe' |
| 'please' | 'censored' | 'blessed' |
| .. | .. | .. |
| 'compament' | 'childcare' | 'braincovid' |

### 2.5 Feature Selection

Feature selection is a process for selecting a subset of features from a set of original features. The use of feature selection is useful for reducing the number of features, speeding up the process of data mining algorithms, and improving data mining performance by eliminating irrelevant, redundant, and noisy features. Irrelevant/appropriate features and redundant features can affect the classification results, so data identification must be carried out [9]. In this study, the feature selection used is information gain. Information gain is a feature selection method that can optimize the performance of a model by sorting features and detecting features that have the most information based on certain classes. So, from the initial features that have as many as 5400 features, the Information gain is selected to be 3000 features. Information gain calculations can be seen in formulas (4) and (5).

$$Entropy_{(S)} = \sum_{i=1}^{c} - p_i \ log_2 \ p_i \quad (4)$$

$$IG_{(S,A)} = Entropy_{(S)} - \sum values_A \frac{|S_v|}{|S|} Entropy_{(S_v)} \quad (5)$$

Where A is an attribute, v is a possible value in attribute A, $values_A$ is the set of possible values for A,

DOI: https://doi.org/10.29207/resti.v7i1.4871

$|S_v|$ is the number of samples for the value v, $|S|$ is the sum of all data samples, and $Entropy(S_v)$ is the entropy value in attribute A.

2.5 KNN Classification

K-Nearest Neighbor is one of the supervised learning algorithms. This algorithm works by classifying a set of data based on training data that has been classified [9], [12]. KNN (K-Nearest Neighbor) can classify based on training data as seen from the closest distance of a data based on the value of k. The distance in question can be calculated using the Euclidean distance equation. The Euclidean distance equation serves to calculate a distance that can interpret the closeness of the distance between two objects. The Euclidean distance equation can be seen in formula (6). Where dist(x,y) represents the distance between the vector and an n-dimensional object.

$$dist(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \quad (6)$$

After getting the Euclidean distance value, then sort the object into the group with the smallest distance. The predicted value will be determined based on the closeness of the training data to the testing data.

2.6 Confusion Matrix

The confusion matrix is one method that can be used to measure the performance of a classification method. The confusion matrix has information obtained by comparing the results of the classification carried out by the system with the results of the classification it should have [15]. The confusion matrix contains actual and predictive information on the classification system. In measuring performance using the confusion matrix, there are four terms to represent the results of the classification, including True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) [15]. Based on the results of these classifications, precision, recall, f1-score, and accuracy values can be obtained.

$$precision = \frac{TP}{TP+FP} \quad (7)$$

$$recall = \frac{TP}{TP+FN} \quad (8)$$

$$\frac{1}{F1} = \frac{1}{2}\left(\frac{1}{precision} + \frac{1}{recall}\right) \quad (9)$$

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (10)$$

The precision value is the ratio of the amount of relevant information obtained by the system to the total amount of information retrieved by the system [16]. The precision value can be obtained using equation (7). The recall value is the ratio of the amount of relevant information obtained by the system to the total amount of relevant information contained in the information, whether or not it is retrieved by the system [16]. The recall value can be obtained using equation (8). The f1-score is the harmonic average between the precision and recall values. The f1-score value can be obtained from equation (9). The accuracy value is the effectiveness of the test based on the effectiveness between the predicted value and the actual value [16]. The accuracy value can be obtained from equation (10).

## 3. Results and Discussions

Tweet crawling data using the Twitter API and Python scripts can be seen in Table 1. The process of crawling tweet data resulted in 179109 tweets or tweets containing the #covid19 hashtag accompanied by user profiles such as username, location, user description, account creation time, number of followers, as well as user verification. However, in this study, only 10000 tweets were used which would then be cleaned through the preprocessing stage, which consisted of case folding, remove punctuation, tokenization, lemmatization, and stop word removal (Table 2).

Clean data obtained from the preprocessing process is given a label of 3000 data. Labeling was carried out in parallel consisting of hoax/valid/neutral labels and credible/not credible labels (Table 4). This is done because this study aims to classify news on Twitter based on author credibility.

The hoax/valid/neutral label is obtained by checking the truth of news from official sites such as the WHO website and news portals such as CNN, Nytimes, Theguardian, Foxnews, and also Huffingtonpost. Credible/non-credible labels are obtained from determinations based on user profiles. If the number of followers is greater than 8000, the number of friends is greater than 10000, the number of likes is greater than 1000, there is a location where the user tweeted, and the year the account was created is under 2022, then the user is considered credible which is represented by number 1 (Table 4).

Table 4. Data Labeling

| text | credibility | hoax |
|---|---|---|
| covid cases still response please cancel compament | 0 | hoax |
| rajasthan government today staed plasma bank sawai singh hospital jaipur treatment covid | 0 | valid |
| nagaland police covid awareness city tower junction dimapur | 0 | valid |
| covid update infection rate florida following natural curve expes predicted initial | 0 | valid |
| coronavirus south africa covid update south africa july | 1 | valid |
| fiji active cases novel coronavirus diseasecovid | 0 | hoax |
| summer garden covid time | 0 | neutral |
| preprint country distancing reveals effectiveness travel restrictions during covid | 0 | hoax |
| rajasthan government today staed | 0 | hoax |

DOI: https://doi.org/10.29207/resti.v7i1.4871

| text | credibility | hoax |
|---|---|---|
| plasma bank sawai singh hospital jaipur treatment covid | | |
| seven years after this talk what glean from education children during covid pandemic | 1 | valid |
| drug treats cough covid ehanire discloses while advising healthy foods with righ | 1 | hoax |
| barcelona legend xavi tests positive covid fans wish midfield maestro speedy recovery tweets | 1 | valid |
| .. | .. | .. |
| herman cain remains oxygen hospitalized covid read | 0 | valid |

After labeling the data, there were 1586 credible data and 1425 non-credible data as well as 2742 valid data, 239 hoax data, and 30 neutral data. Then test the normality of the data to see the distribution of data visualized in the histogram (Figure 2) and create a word cloud to see the frequency of words (Figure 3). It is known in Figure 2, the histogram is in the form of a bell which means that the data is normally distributed.

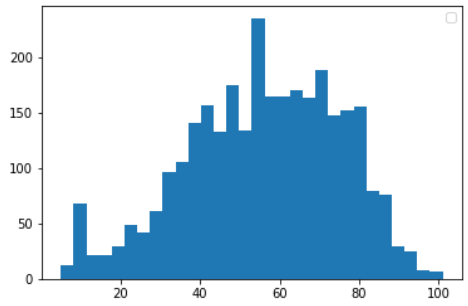


Figure 2. Histogram Data distribution

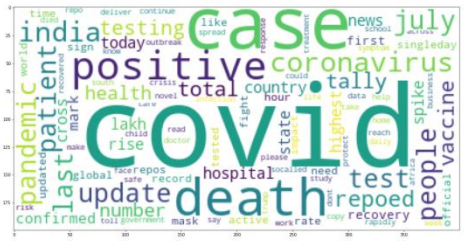


Figure 3. Word Cloud Data

The next step shown ini Table 5 is feature extraction by calculating TF-IDF (Term Frequency Inverse Document Frequency) to see how often the frequency of a word appears in the document [17]. TF-IDF works by converting a document into a matrix per word to calculate its weight. The more often a word appears, the greater the weight of the word.

Table 5. TF-IDF Weight

| words | weight |
|---|---|
| 'covid' | 0.39542906860729055 |
| 'case' | 0.3670425098446454 |
| 'still' | 0.3632854004078306 |
| 'response' | 0.07152390162583482 |
| 'please' | 0.5605575493188689 |
| .. | .. |
| 'compament' | 0.47585752811647847 |

The next process is feature selection using Information Gain. The use of Information Gain in this classification is to speed up the classification process and reduce irrelevant features [11], [18]. Calculation of entropy and Information Gain can be seen in formulas (4) and (5). While the value of Information Gain can be seen in Table 6.

Table 6. Calculation of Information Gain

| words | IG |
|---|---|
| 'covid' | 0.008189305871450087 |
| 'case' | 0.009514473354323483 |
| 'still' | 0.016575774044044644 |
| 'response' | 0.9911667529846511 |
| 'please' | 0.0029943573691557854 |
| .. | .. |
| 'compament' | 0.003543895079013886 |

The division of data into training data and testing data is carried out in 2 scenarios, with the first scenario at 70%:30% and the second scenario at 60%:40%. This is done to develop statistical models and evaluate the performance of machine learning models. In Table 7 the first scenario divides the data into 2053 training data and 880 testing data. Meanwhile, the second scenario divides the data into 1759 training data and 1174 testing data.

Table 7. Split Datasets

| | Scenario I | Scenario II |
|---|---|---|
| Training Data | 2053 | 1749 |
| Testing data | 880 | 1174 |

After the data-sharing process is complete, proceed with the KNN (K-Nearest Neighbor) process. KNN is a learning algorithm that is also known as a category classification algorithm [12], [19]. The classification process using KNN is used to classify documents, where distance and similarity calculations are calculated between the training data and testing data using the Euclidean distance formula. The Euclidean distance formula can be seen in formula (6).

The workflow of the KNN algorithm in this study is choosing the number of K neighbors. Then, calculating the distance from the number of K neighbors using the Euclidean Distance formula. Next, take the K nearest neighbors according to the distance previously calculated using the Euclidean Distance formula. Then, count the number of data points in each category. The final step is to assign a new data point to the category with the most number of neighbors.

Next, in Table 8 a correlation calculation between credibility and hoax is performed. Correlation calculations in this case use Pearson correlation calculations. Pearson correlation is a correlation technique used to determine the relationship between 2 variables (credibility and hoax) [20].

The value of the Pearson correlation for credibility and hoaxes is 0.1148 (scenario I) and 0.1151 (scenario II), meaning that there is a relationship between credibility

DOI: https://doi.org/10.29207/resti.v7i1.4871

and hoax. So, if users who have credibility spread a news story, then the news can be confirmed to be true. In addition, there are also p-values of 0.0006 (scenario I) and 0.0007 (scenario II). Because the p-value <0.05, it can be seen that the data is significant. The Pearson correlation formula can be seen in formula (11). Where $r_{xy}$ represents the Pearson r correlation coefficient, n is the number of samples/observations, x is the independent variable/first variable, and y is the dependent variable/second variable.

Table 8. Correlation Calculation Results

| | Scenario I | Scenario II |
|---|---|---|
| Correlation | 0.1148 | 0.1151 |
| P-value | 0.0006 | 0.0007 |

$$r_{xy} = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{\{n\sum x^2 - (\sum x)^2\}\{n\sum y^2 - (\sum y)^2\}}} \quad (11)$$

The last process in this research is the evaluation or measurement of system performance. System evaluation in this study uses a confusion matrix shown in Figure 4 by calculating classification results that can be predicted correctly or incorrectly using precision, recall, f1-score, and accuracy values. Formulas for calculating precision, recall, f1-score, and accuracy can be seen in formulas (7), (8), (9), and (10).

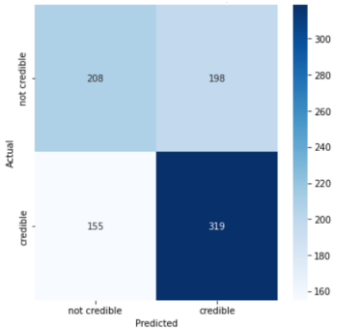


Figure 4. Confusion Matrix Credibility (Scenario I)

Figure 4 displays the credibility confusion matrix for scenario I. Which is known that users are not credible in actual data, it is predicted that as many as 208 users are not credible on the system and it is predicted that as many as 198 are credible on the system. Furthermore, based on actual data, it is predicted that as many as 155 users are not credible on the system and 319 users are credible on the system.

Figure 5 displays the hoax confusion matrix for scenario I. It is known that hoaxes in actual data are predicted as many as 8 data as hoaxes in the system, 63 data are predicted as valid in the system, and 0 data are predicted as neutral in the system. Furthermore, valid on actual data, predicted as many as 6 data as hoaxes on the system, predicted 796 data as valid on the system, and predicted as many as 0 data as neutral on the system. After that, neutral on actual data, predicted as many as 0 data as hoaxes on the system, predicted 7 data as valid on the system, and predicted as many as 0 data as neutral on the system.

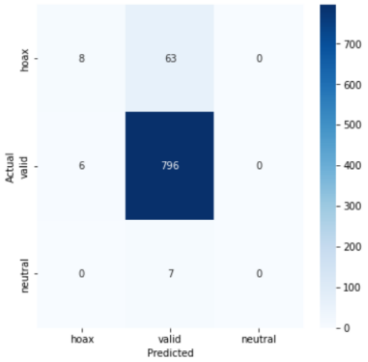


Figure 5. Confusion Matrix Hoax (Scenario I)

Table 9. Value of Confusion Matrix Credibility (Scenario I)

| | Precision | Recall | F1-score |
|---|---|---|---|
| 0 | 0.57 | 0.51 | 0.54 |
| 1 | 0.62 | 0.67 | 0.64 |
| accuracy | | | 0.60 |

Table 10. Value of Confusion Matrix Hoax (Scenario I)

| | Precision | Recall | F1-score |
|---|---|---|---|
| 1 | 0.57 | 0.11 | 0.19 |
| 2 | 0.92 | 0.99 | 0.95 |
| 3 | 0.00 | 0.00 | 0.00 |
| accuracy | | | 0.91 |

The confusion matrix produces information that compares the predicted results of the classification performed by the system with the actual classification. Table 9 and 10 show the confusion matrix, there are 4 terms to represent the results of the classification, namely TP (True Positive), TN (True Negative), FP (False Positive), and FN (False Negative). TP (True Positive), namely the amount of positive data that is classified as true by the system, TN (True Negative), namely the amount of negative data that is classified as true by the system, FP (False Positive), namely the amount of positive data that is classified as wrong by the system, and FN (False Negative) ) is the amount of negative data that is classified incorrectly by the system.

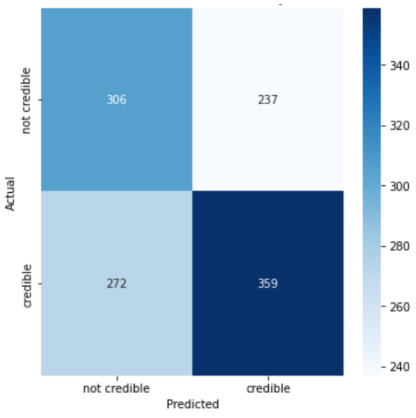


Figure 6. Confusion Matrix Credibility (Scenario II)

DOI: https://doi.org/10.29207/resti.v7i1.4871

Figure 6 displays the credibility confusion matrix for scenario II. Which is known that users are not credible on actual data, it is predicted that as many as 306 users are not credible on the system and it is predicted that as many as 237 are credible on the system. Furthermore, based on actual data, it is predicted that as many as 272 users are not credible on the system and 359 users are credible on the system.

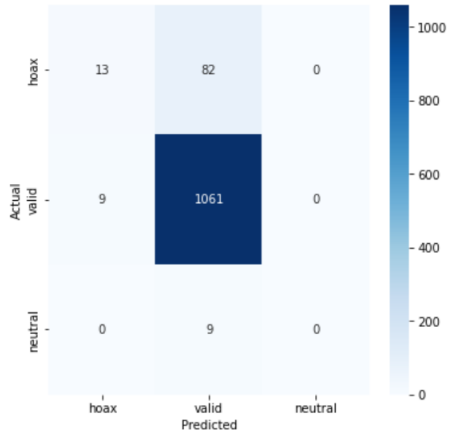


Figure 7. Confusion Matrix Hoax (Scenario II)

Figure 7 displays the hoax confusion matrix for scenario II. Which is known that hoaxes in actual data, predicted as many as 13 data as hoaxes in the system, predicted 82 data as valid in the system, and predicted as many as 0 data as neutral in the system. Furthermore, valid on actual data, predicted as many as 9 data as hoaxes on the system, predicted 1061 data as valid on the system, and predicted as many as 0 data as neutral on the system. After that, neutral on actual data, predicted as many as 0 data as hoaxes on the system, predicted 9 data as valid on the system, and predicted as many as 0 data as neutral on the system.

Table 11. Value of Confusion Matrix Credibility (Scenario II)

| | Precision | Recall | F1-score |
|---|---|---|---|
| 0 | 0.53 | 0.56 | 0.55 |
| 1 | 0.60 | 0.57 | 0.59 |
| accuracy | | | 0.57 |

Table 12. Value of Confusion Matrix Hoax (Scenario II)

| | Precision | Recall | F1-score |
|---|---|---|---|
| 1 | 0.59 | 0.14 | 0.22 |
| 2 | 0.92 | 0.99 | 0.95 |
| 3 | 0.00 | 0.00 | 0.00 |
| accuracy | | | 0.91 |

Tables 11 and 12 show the credibility and hoax confusion matrix values in scenario II. Where the precision value is obtained from a comparison of the amount of relevant information obtained by the system with the total amount of information retrieved by the system. Furthermore, the recall value is obtained from a comparison of the amount of relevant information obtained by the system with the total amount of relevant information contained in the information, whether retrieved or not retrieved by the system. Then, the f1-score value is obtained from the average harmonic result between the precision and recall values. Meanwhile, the accuracy value indicates the effectiveness of the test based on the effectiveness between the predicted value and the actual value.

## 4. Conclusion

Based on the results and discussion, it can be concluded that the detection of fake news with tweets containing the #covid19 hashtag based on author credibility using the Information Gain and KNN (K-Nearest Neighbor) methods was successfully carried out with an accuracy value of 91%, a correlation value between credibility and hoaxes of 0.115, and a p-value <0.05. This proves that the system is 91% accurate and there is a significant correlation between credibility and hoaxes. Scenario division into 2 scenarios also has a significant impact on precision, recall, and f1-score. Where in the scenario I, precision is 0.60, recall is 0.59, and f1-score is 0.59 for credibility. And precision is worth 0.50, recall is worth 0.37, and f1-score is worth 0.38 for hoaxes. Meanwhile, in scenario II, precision is 0.57, recall is 0.57, and f1-score is 0.57 for credibility. And precision is worth 0.50, recall is worth 0.38, and f1-score is worth 0.39 for hoaxes.

For research development, it is necessary to do other news classifications, not only COVID-19 in Indonesia. In addition, it is necessary to research whether the calculation of the confusion matrix between credibility and hoaxes can be combined. In addition, the addition of using feature selection (not only Information Gain) and using other classification methods to improve system performance.

## References

[1] K. K. Kapoor, K. Tamilmani, N. P. Rana, P. Patil, Y. K. Dwivedi, and S. Nerur, "Advances in Social Media Research: Past, Present, and Future," *Inf. Syst. Front.*, vol. 20, no. 3, pp. 531–558, Jun. 2018, doi: 10.1007/s10796-017-9810-y.

[2] N. S. Mudawamah, "Internet User Behavior: Case Study of Library and Information Science Department Students of UIN Maulana Malik Ibrahim".

[3] S. Alhabash and M. Ma, "A Tale of Four Platforms: Motivations and Uses of Facebook, Twitter, Instagram, and Snapchat Among College Students?," *Soc. Media Soc.*, vol. 3, no. 1, p. 205630511769154, Jan. 2017, doi: 10.1177/2056305117691544.

[4] G. K. Shahi, A. Dirkson, and T. A. Majchrzak, "An Exploratory Study of Covid-19 Misinformation on Twitter," *Online Soc. Netw. Media*, vol. 22, p. 100104, Mar. 2021, doi: 10.1016/j.osnem.2020.100104.

[5] L. Garrett, "Covid-19: the Medium is the Message," *The Lancet*, vol. 395, no. 10228, pp. 942–943, Mar. 2020, doi: 10.1016/S0140-6736(20)30600-0.

[6] J. Zarocostas, "How to Fight an Infodemic," *The Lancet*, vol. 395, no. 10225, p. 676, Feb. 2020, doi: 10.1016/S0140-6736(20)30461-X.

[7] J. Zhang, B. Dong, and P. S. Yu, "Fake Detector: Effective Fake News Detection with Deep Diffusive Neural Network."

DOI: https://doi.org/10.29207/resti.v7i1.4871

arXiv, Aug. 10, 2019. Accessed: Jan. 14, 2023. [Online]. Available: http://arxiv.org/abs/1805.08751

[8] T. R. A. Pangaribuan, "Social Media Credibility in Reporting the Jakarta Governor Election," vol. 18, no. 2.

[9] R. I. Pristiyanti, M. A. Fauzi, and L. Muflikhah, "Sentiment Analysis Summarizing Film Reviews Using Information Gain and K-Nearest Neighbor Methods".

[10] I. Maulida, A. Suyatno, and H. R. Hatta, "Feature Selection in Abstract Indonesian Text Documents Using the Information Gain Method," *J. SIFO Mikroskil*, vol. 17, no. 2, pp. 249–258, Oct. 2016, doi: 10.55601/jsm.v17i2.379.

[11] R. K. Dinata, H. Novriando, N. Hasdyna, and S. Retno, "Attribute Reduction Using Information Gain for Cluster Optimization of the K-Means Algorithm," *J. Edukasi Dan Penelit. Inform. JEPIN*, vol. 6, no. 1, p. 48, Apr. 2020, doi: 10.26418/jp.v6i1.37606.

[12] R. I. Perwira, B. Yuwono, R. I. P. Siswoyo, F. Liantoni, and H. Himawan, "Effect of Information Gain on Document Classification Using K-Nearest Neighbor," *Regist. J. Ilm. Teknol. Sist. Inf.*, vol. 8, no. 1, p. 50, Jan. 2022, doi: 10.26594/register.v8i1.2397.

[13] S. Tang, S. Yuan, and Y. Zhu, "Data Preprocessing Techniques in Convolutional Neural Network Based on Fault Diagnosis Towards Rotating Machinery," *IEEE Access*, vol. 8, pp. 149487–149496, 2020, doi: 10.1109/ACCESS.2020.3012182.

[14] R. Ahuja, A. Chug, S. Kohli, S. Gupta, and P. Ahuja, "The Impact of Features Extraction on the Sentiment Analysis," *Procedia Comput. Sci.*, vol. 152, pp. 341–348, 2019, doi: 10.1016/j.procs.2019.05.008.

[15] O. Caelen, "A Bayesian Interpretation of the Confusion Matrix," *Ann. Math. Artif. Intell.*, vol. 81, no. 3–4, pp. 429–450, Dec. 2017, doi: 10.1007/s10472-017-9564-8.

[16] Isman, Andani Ahmad, and Abdul Latief, "Comparison of KNN and LBPH Methods in Classification of Herbal Leaves," *J. RESTI Rekayasa Sist. Dan Teknol. Inf.*, vol. 5, no. 3, pp. 557–564, Jun. 2021, doi: 10.29207/resti.v5i3.3006.

[17] G. B. Firmanesha, S. S. Prasetyowati, and Y. Sibaroni, "Detecting Hoax News Regarding the Covid-19 Vaccine Using Levenshtein Distance".

[18] A. Essra, "Analysis of Information Gain Attribute Evaluation for Classification of Intrusion Attacks," 2016.

[19] M. I. A. Ismandiya and Y. Sibaroni, "Indonesian News Classification Using Weighted K-Nearest Neighbour".

[20] I. J. A. Cici Apriza Yanti, "Pearson, Spearman, and Kendall Tau Correlation Test Differences in Analyzing the Incidence of Diarrhea," *J. Endur.*, vol. 6, no. 1, pp. 51–58, Jun. 2022, doi: 10.22216/jen.v6i1.137.

.

DOI: https://doi.org/10.29207/resti.v7i1.4871