Bit-Tech

Vol.4, No.2, December 2021
Available online at: http://jurnal.kdi.or.id/index.php/bt

Implementation of Naïve Bayes Algorithm for Classification of Mental Health of Social Media Users

Aditiya Hermawan[1)*]

[1)] *Universitas Buddhi Dharma*
Jl. Imam Bonjol No.41, Tangerang, Indonesia
[1)]aditiya.hermawan@ubd.ac.id

Article history:

Received 12 December 2021;
Revised 15 December 2021;
Accepted 23 December 2021;
Available online 30 December 2021

Keywords: {use 4-6 keywords}

Naïve Bayes
Classification
Mental Health
Social Media

Abstract

Social media has become a human need to interact in everyday life. Apart from being a means of communication, social media also has the additional function of exchanging information on the internet in various forms including writing, images, and videos. One of the social media that has many users is Instagram, where Instagram offers information sharing features in the form of images, photos, and short videos. The purpose of this feature is for users to express themselves and attract the attention of others, thereby creating feelings of happiness and increasing self-confidence. In addition to positive impacts, there are also negative impacts on users, for example excessive use that causes addiction so that it can cause mental health disorders. Mental health needs to be handled properly so that it does not continue to get worse, but there are several obstacles in seeing a psychiatrist in mental health, including limited access and negative stigma if someone sees a psychiatrist. Therefore, a tool is needed that can be an early indication in knowing the level of mental agitation, especially in the use of Instagram. Classification in data mining can help provide initial information on a person's condition in his mental health. The Naïve Bayes algorithm provides an accuracy rate of 92.5% in classifying mental health on data sets that have been clustered. Good accuracy can help social media users know their mental health condition.

I. INTRODUCTION

The use of social media in everyday life has become one of the necessities for humans to interact. Social media itself has several functions, ranging from means of communication and as a medium for exchanging information on the internet, starting from information in the form of writing, images, and videos. One of the most popular social media is Instagram with active users worldwide reaching 928,500,000 users as of January 2020[1]. Instagram is a social media that can share information in the form of images or photos and short videos. Instagram also provides inspiration for its users and can increase one's creativity, because it has features that can make photos more beautiful, artistic, and good [2]. The purpose of this feature is for users to be able to express themselves and attract attention from others, thereby creating feelings of happiness, worth and increased self-confidence due to getting attention from others.

In addition to the positive side of Instagram, there are also negative impacts on its users. For example, users who use Instagram more than 2 hours a day and 58 times a week are more prone to experiencing adverse effects on their users [3]. Some of the bad effects that arise such as anxiety, depression, intimidation, loneliness, lack of sleep and impaired self-image and fear of not getting information about others. These adverse effects can make a person's mental health unhealthy. Therefore, Instagram is one of the social media that has a bad impact on the mental health of its users [4].

Mental health is a condition of a person who can realize his own potential and can also cope with stress and can work productively and contribute positively to society at large [5]. Determining a person's mental health is something that can be seen directly, but must take into account several indicators. If a person experiences a mental disorder or an unhealthy mental condition, a psychologist or psychiatrist should be able to help a person recover mentally. However, there are several obstacles in meeting the psychologist, ranging from limited access to psychologists and also

* Corresponding author

ISSN 2622-2728 (online) 2622-271X (print)
doi:

awareness about the importance of mental health, not to mention the stigma that thinks if someone goes to a psychologist or psychiatrist, it indicates that the person has a mental disorder.

Instagram users, especially in Indonesia, which ranks fourth with the most users, amounting to 63 million active users in the first quarter of 2020, it is reported that 70% of users feel mental disorders due to using Instagram [6]. Based on the impact that can be caused by social media in this case is Instagram and the low level of public awareness about the importance of knowing mental health due to the use of Instagram, a tool is needed that can measure the level of mental health of Instagram users as an initial indication.

To see the level of mental health, accurate information about mental health conditions is needed with several predetermined indicators and data that has been tested. The data used is the result of clustering Instagram users which has been validated by psychologists. To be able to see if someone has a mental health disorder or not, we need a model that can classify and have a good level of accuracy from existing data using the help of Data Mining. Data mining is a technique that finds new patterns from a large data set and includes methods that are slices of artificial intelligence, machine learning, statistics, and database systems [7]. There is a role for data mining, namely, to classify data and provide information on its level of accuracy. One of the algorithms that can be used for classification is nave Bayes, Nave Bayes has the ability comparable to decision trees and neural networks to make data classifications [8].

II. RELATED WORKS/LITERATURE REVIEW

A. Data Mining

Data mining is a combination of several disciplines of computer science that finds new patterns from a large data set and includes methods that are slices of artificial intelligence, machine learning, statistics, and database systems [7]. The results of data processing using data mining methods are useful for making decisions in the future. In general, data mining discusses methods such as classification, clustering, association, regression. Classification is the process of finding a model (function) that describes and distinguishes data classes or concepts that aim to be used to predict the class of objects whose class label is unknown [1]. Data classification consists of 2 process steps. The first is learning (training phase), where a classification algorithm is created to analyze the training data and then it is represented in the form of a classification rule. The second process is classification, where test data is used to estimate the accuracy of the classification rule [8].

B. Naïve Bayes

Naïve Bayes, which is also called idiot's Bayes, simple Bayes, and independence Bayes, is a good method because it is easy to make, does not require a complicated iteration parameter estimation scheme, which means it can be applied to large data sets [9]. The use of the Bayes theorem in the Naïve Bayes algorithm is to combine prior probabilities and conditional probabilities in a formula that can be used to calculate the probability of each possible classification [10].

The stages in the Naive Bayes method are:

1. Determination of data categories

   Naive Bayes classifies the probability values of p (C=ci | D=dj), namely the probability of ci and the data of dj. Classification is done to determine c C from data d D where C= {c1, c2, c3, ..., ci} and D= {d1, d2, d3, d... dj}. Determination of the category of a data is done by finding the maximum value of p (C=ci | D=dj) at P= {p (C=ci | D=dj) | c C and d D}. The probability value of p (C=ci | D=dj) can be calculated by Mitchell's 2005 equation:

$$p(C = ci \mid D = dj) = \frac{P(C = ci \cap D = dj)}{P(D = dj)} = \frac{p(D = dj \mid C = ci)\, x\, p(C = ci)}{p(D = dj)} \quad \text{....(1)}$$

   Where p (D=dj | C=ci) is the probability value of the occurrence of the dj document if it is known that the data is in the ci category, p(C=ci) is the probability value of the occurrence of the ci category and p (D= dj) is the probability value of the occurrence of the dj document.

2. Calculating probability

Naive Bayes considers a data as a collection of words that compose the data and does not pay attention to the order in which the words appear in the data. So, the calculation of the probability p (D=dj | C=ci) can be considered as the product of the multiplication of the probability of occurrence of words in the data dj. The calculation of the probability p (C=ci | D=dj) can be written as follows:

$$p(C = ci \mid D = dj) = \frac{\prod_k p(wk \mid C = ci)\ x\ p(C = ci)}{p(w1, w2, w3, \dots, wk, \dots wn)} \quad \text{.....(2)}$$

where $\prod_k p(wk \mid C = ci)$ is the product of the probability occurrences of all words in the dj data?

3. Classification process
   The classification process is carried out by making a probabilistic model from the training data, namely by calculating the value of p (wk | c). For discrete wkj with wkj V = {v1, v2, v3, ..., vm} then the value of p (wk | c) is searched for all possible values of wkj and obtained by doing the calculation:

$$p(wk = wkj \mid C) = \frac{Db(wk = wkj, c)}{Db(c)} \quad \text{dan } p(c) = Db(c)/|D|$$

   Where Db (wk = wkj, c) is a function that returns the number of documents b in category c which has the word value wk = wkj, Db(c) is a function that returns the number of documents b that have categories c and |D| is the total number of training data. The equation Db (wk = wkj, c) is often combined with Laplacion Smoothing to prevent the equation Db (wk = wkj, c) from being written as:

$$p(wk = wkj \mid C) = \frac{Db(wk = wkj, c) + 1}{Db(c) + |V|} \quad \text{.....(3)}$$

4. Categorizing values
   With |V| is the sum of the possible values of wkj. The categorization of a data is done by selecting the value of c which has the maximum p (C=ci | D=dj) value, and is expressed by:

$$c* = \arg \max_{c \in C} p \prod_k p(wk \mid c)\ x\ p(c) \quad \text{.....(4)}$$

   Category c* is a category that has a maximum p (C=ci | D=dj) value. The value of p(D=dj) does not affect the comparison because for each category the value will be the same. P for each variable can be formulated as follows:

$$P(x = v|c) = \frac{1}{\sqrt{2\pi\sigma_c^2}}\, e^{-\frac{(v-\mu_c)^2}{2\sigma_c^2}} \quad \text{.....(5)}$$

C. Mental Health

Mental health is an individual who is free from all psychiatric symptoms or mental disorders with the realization of harmony between mental functions and can deal with problems that occur [11]. Mental health is the realization of true harmony between mental functions and the creation of adjustment between humans and themselves and their environment based on faith and piety and aims to achieve a meaningful and happy life in this world and the afterlife.[2].

Problems in mental health are problems that must be addressed comprehensively and there are several factors that influence it, namely as follows: [3]:

a. Physical factor (*organobiologis*)
   Physical factors can affect the quality of a person's mental health, for example, when a person finds out that his body is being devoured by cancer and at that moment the person has lost part of his life, even though he is consciously aware that his mental and emotional state has been disturbed which accelerates the process of decreasing the immune system. drastically and his spirit of life also decreased.
b. Emotional factor (*psikoedukatif*)
   Strength in mental, emotional, and positive suggestions that support is needed to awaken the spirit of life in restoring health both physically and spiritually.

c. Socio-cultural factors (*sosiokultural*)

In the environment where you live, family and culture greatly determine the quality of a person's mental and emotional health, and communication between people is very much needed in dealing with all the problems that come in life.

III. METHODS

The method used for classification is the Knowledge Discovery Database (KDD) which is the core of the Data Mining process, where the algorithm explores data, builds models, and finds unknown patterns. KDD is a problem solving by analysing the existing data in the database. KDD has process stages as in Fig 1 [4].

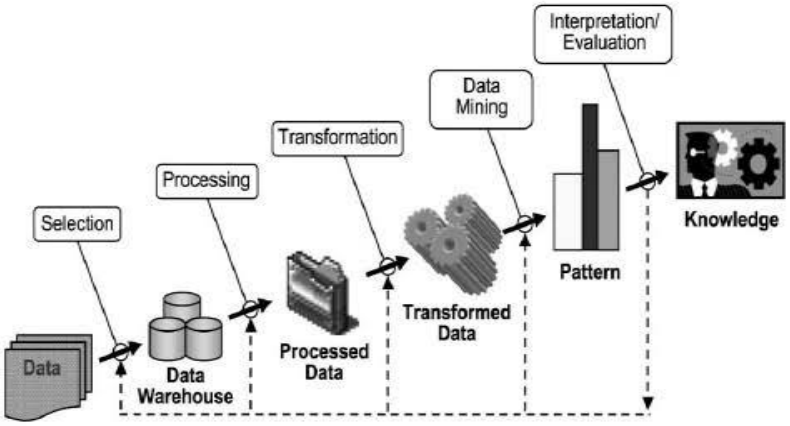

Fig. 1 KDD Process Stages [4]

1. Data Selection
   Data selection needs to be done before the stage of extracting information in KDD begins. The results of the selection of data to be used for the data mining process are stored in one file and separated from the operational database.
2. Pre-processing / Cleaning
   This cleaning process includes removing duplication of data, checking for inconsistent data, correcting errors in data, carrying out an enrichment process or enriching existing data with relevant and required data in KDD.
3. Transformation
   The transformation process on the data that has been selected, so that the data is in accordance with the data mining process. This is a creative process and is highly dependent on the type or pattern of information to be searched in the database
4. Data Mining
   Data mining is the process of looking for interesting patterns or information in selected data using certain techniques or methods in accordance with the overall goals and processes of KDD.
5. Interpretation / Evaluation
   This stage includes examining the information generated against existing facts or hypotheses and displaying the information obtained from the data mining process in a form that is easily understood by users.

IV. RESULTS

A. Data Selection

The data used are primary data obtained using a questionnaire. The questionnaire is intended for 600 Instagram users in Indonesia. The following is the number of respondents by gender and age:

TABLE 1
RESPONDENT

Usia	Jenis Kelamin	
	Laki-laki	Perempuan
< 18	12	13
18 - 25	124	180
26 - 45	108	142
> 45	10	11
Total	254	346

In obtaining the data, the selection or selection of data is carried out from a set of data obtained from the questionnaire and carried out before the information mining stage. In the selection of the data, the attributes that will be used in this study have been selected. The attributes used are as follows:

TABLE 2
DETAIL ATRIBUT [5]

No.	Atribut	Keterangan
1	Acc	Number of instagram accounts
2	Freq	Frequency of accessing Instagram in a day
3	Durasi	Duration of accessing Instagram in a day
4	Candu	Addicted to access
5	Edit	Photo/video editing
6	Like	Likes for photos/videos
7	Com	Comment
8	Iri	Envy other people's photos/videos
9	FOMO	Fear of missing information
10	Foll	Account followers

B. Data Cleaning

At this stage, deletion of data that is empty, remove duplication and check inconsistent data, correct errors in data, and delete some attributes that are not used in this study.

TABLE 3
SAMPLE DATA SET

No	Nama	Usia	JK	Acc	Freq	Durasi	Candu	Edit	Like	Com	Iri	FOMO	Foll
1	AA	45	P	1	less than 8 times/day	less than 1 hour/day	2	1	4	1	1	1	5
2	AA	16	L	2	8-16 times/day	more than 2 hours/day	4	5	3	2	1	2	4
3	AdP	16	L	1	8-16 times/day	more than 2 hours/day	3	1	1	1	1	1	1
4	ADS	20	P	1	less than 8 times/day	less than 1 hour/day	2	1	1	1	2	1	1
5	AE	15	P	2	more than 16 times/day	more than 2 hours/day	5	2	3	1	1	3	5
6	AF	27	P	1	8-16 times/day	less than 1 hour/day	3	3	3	2	2	2	4
7	AG	27	P	1	8-16 times/day	more than 2 hours/day	5	5	3	4	3	4	4
8	Ag	19	P	2	more than 16 times/day	less than 1 hour/day	4	4	2	2	1	2	5
9	AH	50	L	3	8-16 times/day	1-2 hours/day	5	5	5	5	1	5	5
10	AH	50	L	3	8-16 times/day	1-2 hours/day	5	5	5	5	1	5	5

At this stage, deletion of several attributes that are not used in the classification process such as name, age, gender, and removing duplication of data, as well as correcting errors in the data.

TABLE 4
DATA SET AFTER SELECTION

No	Acc	Freq	Durasi	Candu	Edit	Like	Com	Iri	FOMO	Foll
1	1	less than 8 times/day	less than 1 hour/day	2	1	4	1	1	1	5
2	2	8-16 times/day	more than 2 hours/day	4	5	3	2	1	2	4
3	1	8-16 times/day	more than 2 hours/day	3	1	1	1	1	1	1
4	1	less than 8 times/day	less than 1 hour/day	2	1	1	1	2	1	1
5	2	more than 16 times/day	more than 2 hours/day	5	2	3	1	1	3	5
6	1	8-16 times/day	less than 1 hour/day	3	3	3	2	2	2	4
7	1	8-16 times/day	more than 2 hours/day	5	5	3	4	3	4	4
8	2	more than 16 times/day	less than 1 hour/day	4	4	2	2	1	2	5
9	3	8-16 times/day	1-2 hours/day	5	5	5	5	1	5	5
10	3	8-16 times/day	1-2 hours/day	5	5	5	5	1	5	5

C. Data Transfromation

At this stage, discretization is carried out to match the expected analysis and the required data format, to facilitate the data mining process. Discretization is carried out on the following attributes:

1. Frequency:
   a. Less than 8 times/day = 1
   b. 8-16 times/day = 2
   c. More than 16 times/day = 3
2. Duration:
   a. Less than 1 hour/day = 1
   b. 1-2 hours/day = 2
   c. More than 2 hours/day = 3

TABLE 5
DATA AFTER TRANSFORMATION

No	Acc	Freq	Durasi	Candu	Edit	Like	Com	Iri	FOMO	Foll
1	1	1	1	2	1	4	1	1	1	5
2	2	2	3	4	5	3	2	1	2	4
3	1	2	3	3	1	1	1	1	1	1
4	1	1	1	2	1	1	1	2	1	1
5	2	3	3	5	2	3	1	1	3	5
6	1	2	1	3	3	3	2	2	2	4
7	2	3	1	4	4	2	2	1	2	5
8	1	2	3	5	5	3	4	3	4	4
9	1	1	1	3	3	3	4	4	4	3
10	3	2	2	5	5	5	5	1	5	5

D. Data Mining

1. Labeling

From the dataset above, there is still no label or dependent variable that becomes a reference for training in modeling. Labels are obtained through the Clustering method first where the clusters formed are 3 categories of mental health, namely:

1. Healthy,
2. Worrying,
3. Very worrying.

The following is a labeled dataset that will be used in making the classification model.

TABLE 6
DATA DENGAN TABEL

No	Acc	Freq	Durasi	Candu	Edit	Like	Com	Iri	FOMO	Foll	Kesehatan Mental (label)
1	3	2	3	5	5	3	1	1	1	3	Very worrying
2	2	2	3	4	3	2	2	2	3	5	Very worrying
3	1	2	2	3	1	1	2	1	1	3	Healthy
4	1	2	2	2	1	2	2	1	3	3	Healthy
5	2	2	2	3	2	1	2	1	1	4	Healthy
6	1	3	3	4	4	3	3	2	2	4	Very worrying
7	1	2	1	5	4	3	3	3	2	5	Very worrying
8	1	2	1	3	1	3	4	2	1	3	Healthy
9	1	2	2	3	4	1	1	1	1	2	Healthy
10	1	3	3	4	3	2	4	1	2	3	Very worrying

2. Model Design

The design of the model begins with the process of separating the dataset into 2 types of data, namely training data and testing data by dividing the percentage of the total dataset. If the training data used is 10% of the dataset, the percentage of data testing is 90%. And if the training data used is 20% of the dataset, the percentage of data testing is 80%. The amount of training data when added to the testing data will be the number of the dataset. The percentage of the training data will be determined based on the iteration of the cross-validation process. For example, in the cross-validation process in the 4th iteration of a total of 10 iterations, the percentage of training data is 40% and the percentage of testing data is 60%. The process of separating training data and testing data follows a cross-validation process so that the data used in the process continues to use the same dataset. In the process of separating the training data and testing data, we still use random data selection from the dataset.

After separating the datasets, the Training data will be modeled using the Naïve Bayes algorithm, and evaluation using the Testing dataset will be carried out by measuring the accuracy, model performance (f-measure, precision, recall) using a confusion matrix. The Process Flow is depicted as Fig. 3

The following are the results of calculations from existing datasets using Naïve Bayes:

TABLE 7
NAVE BAYES MODEL

Attribute	Parameter	Very worrying	Healthy	Worrying
acc	mean	1.747	1.268	1.23
acc	standard deviation	0.849	0.565	0.502
freq	mean	2.343	1.662	2.183
freq	standard deviation	0.705	0.715	0.675
durasi	mean	2.404	1.515	1.764
durasi	standard deviation	0.724	0.691	0.712
candu	mean	3.831	2.615	3.0
candu	standard deviation	1.076	1.136	0.946
edit	mean	3.803	2.385	2.853
edit	standard deviation	0.992	1.147	0.864
lke	mean	3.427	1.563	1.743
lke	standard deviation	1.098	0.820	0.847
com	mean	3.213	1.762	2.654
com	standard deviation	1.280	0.951	1.199
iri	mean	2.596	1.338	3.168
iri	standard deviation	1.338	0.638	1.250
fomo	mean	2.904	1.472	3.639
fomo	standard deviation	1.265	0.743	0.984
foll	mean	4.129	2.823	3.529

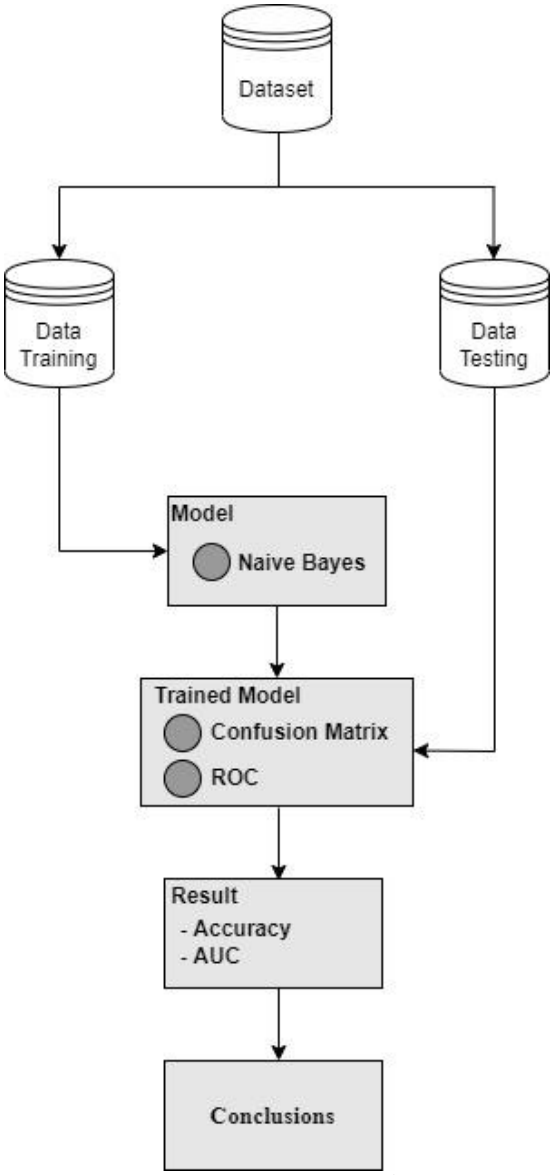

Fig. 2 Mental Health Data Mining Classification Model

E. Evaluation

After finishing making the model with nave Bayes, the next step is to see the accuracy results with the confusion matrix (numbers containing False Positive, False Negative, True Positive, and True Negative).

1. Confusion matrix

   This confusion matrix method uses a matrix table as in Table 8, if the data set only consists of two classes, one class is considered positive and the other negative. [10]

TABLE 8
MODEL CONFUSION MATRIX [10]

Correct classification	**Clasified as**	
	+	-
+	True Positives	False Negatives
-	False Positives	True Negatives

True positives (TP) are the number of positive records classified as positive, false positives (FP) is the number of negative records classified as positive, false negatives (FN) is the number of positive records classified as negative, true negatives (TN) is the number of records negative which is classified as negative. To find the value of Sensitivity, Specificity, Precision and Accuracy, the following equation is used:

Sensitivity $= \frac{TP}{P}$

Specificity $= \frac{TN}{P}$

Precision $= \frac{TP}{TP+FP}$

Accuracy = sensitivity $\frac{P}{P+N}$ + specificity $\frac{N}{P+N}$

Description:

TP = number of true positives
TN = number of true negatives
P = number of positive records
N = number of negative records
FP = number of false positives

TABLE 9
EVALUATION RESULT

	true Very worrying	**true Healthy**	**true worrying**	**class precision**
pred. Sangat Mengkhawatirkan	164	7	12	**89.62%**
pred. Sehat	3	212	0	**98.60%**
pred. Mengkhawatirkan	11	12	179	**88.61%**
class recall	**92.13%**	**91.77%**	**93.72%**	

The accuracy value of the confusion matrix is as follows:

$$akurasi = \frac{(TN + TP)}{(TN + FN + TP + FP)}$$
$$= \frac{(164 + 212 + 179)}{\big((164 + (7 + 12)) + (212 + (3 + 0)) + (179(11 + 12))\big)}$$
$$= \frac{555}{600}$$
$$= 0{,}925 = \mathbf{92,5}\%$$

V. CONCLUSIONS

Based on the accuracy value obtained, which is 92.5%, the classification model made with the Naïve Bayes algorithm can be categorized as Very Good for classifying mental health on social media users so that it can be used as a reference as an initial indication of measuring the mental health of a social media user, especially Instagram.

There are several things that might be investigated further, namely the amount of data that is more, and the validation of clustering formed with more experts so that the dataset that is made the classification model has a higher level of validity, also can be made a classification model with other classification algorithms to get even better results.

VI. REFERENCES

[1] Y. Kurnia, Y. Isharianto, Y. C. Giap, A. Hermawan, and Riki, "Study of application of data mining market basket analysis for knowing sales pattern (association of items) at the O! Fish restaurant using apriori algorithm," in *Journal of Physics: Conference Series*, 2019, vol. 1175, no. 1, doi: 10.1088/1742-6596/1175/1/012047.

[2] Hasneli, *Kesehatan Mental dalam Pandangan Islam*. Padang: Haifa Press, 2014.

[3] A. Hakim, *Hipnoterapi : Cara Tepat dan Cepat Mengatasi Stres, Fobia, Trauma, dan Gangguan Mental Lainnya*. Jakarta: Visimedia, 2010.

[4] R. T. Vulandari, "Pengelompokan Tingkat Keamanan Wilayah Jawa Tengah Berdasarkan Indeks Kejahatan dan Jumlah Pos Keamanan dengan Metode Klastering K-Means," *J. Ilm. SINUS*, vol. 14, no. 2, 2016.

[5] N. A. Mondoano, R. Mayasari, and F. Gunawan, "Instagram dan Kesehatan Mental pada Mahasiswa Institut Agama Islam Negeri Kendari," *ANALITIKA*, vol. 10, no. 2, p. 104, Dec. 2018, doi:

10.31289/analitika.v10i2.2034.