



Classification of Device Addiction to Students using SAS-SV with K-Nearest Neighbor Algorithm

Basyir Al Musthoqfirin Majid*, Abdul Mubarak, Salkin Lutfi

Department of Informatic Engineering, Universitas Khairun Ternate, Indonesia

Correspondence: E-mail: basyialmajid675@gmail.com

ABSTRACT

A gadget is a small electronic device with a particular purpose, often thought of as an innovation of new goods. Not only to help facilitate human activities, but gadgets are also a part of the lifestyle for modern citizens. With this innovative feature, the gadget has attracted users more and more, or in other words, users have become more addicted to the gadget. This study aims to investigate how addictive gadgets are to students at the Department of Informatic Engineering, Khairun University, Ternate, Indonesia using K-Nearest Neighbor (KNN) Algorithm. In KNN, there is a Training dataset where one set of data contains the class's value and a predictor that will be used as one of the requirements for determining a suitable grade per the predictor. In contrast, the Testing dataset contains the new data that will be classified based on the model made and the accuracy of classification in the data collection process. Questionnaires were made using Google forms, then distributed through the internal groups of the Informatics Engineering department of Khairun University. A total of 78 questionnaires were successfully collected. The results showed that the testing accuracy with $k = 3$ is 86% and $k = 5$ is 80%. This show that KNN algorithm can be applied to measure the level of addiction to students.

ARTICLE INFO

Article History:

Submitted/Received 13 Feb 2022

First Revised 25 Feb 2022

Accepted 20 Mar 2022

First Available online 01 Apr 2022

Publication Date 01 Apr 2022

Keyword:

Classification,

Device Addiction,

Gadget,

K-Nearest Neighbor,

SAS-SV.

1. INTRODUCTION

One of the most proliferating technology developments is a gadget. A gadget is a small electronic device with a particular purpose, often thought of as an innovation of new goods. Not only to help facilitate human activities, but gadgets are also a part of the lifestyle for modern citizens. Today, almost everyone owns it, ranging from urban residents to rural residents [1]. In line with the growth of technology, gadgets are now equipped with various applications that can be adjusted based on user's preferences and wills, such as smartphones. Smartphones provide various features such as camera, MP3, GPS, website browsing, calling, e-mail, gaming, and social networking services (SNS) which are incorporated in one portable device.[2] With these innovative features, the gadget has attracted users more and more, or in other words, users have become more addicted to the gadget. Addiction can be interpreted as a condition where people feel dependent on something they enjoy on various opportunities that exist due to the lack of control over attitudes so they feel punished if they do not fulfill their desires and habits [3]. Thus, our goal is to investigate the level of addiction in using gadgets for students at the Department of Informatic Engineering, Khairun University, Ternate, Indonesia using K-Nearest Neighbor (KNN) Algorithm.

KNN is a machine learning algorithm which has been widely used for classification problems by taking the closest distance from the majority of existing neighbor data (training-set). Previous research from Sahambangung, et al used KNN to classify drug users based on the symptoms entered by users. Whereas Satrian et al [4] used KNN to classify the school-age Gamers at the Holy Kalam Christian School. Their study also investigate which online games has an impact on the daily life and academic achievement of students. Their study successfully helps teachers at schools guide students who are addicted to games so that they can change for the better.[5]

Based on those two studies, we aim to apply the K-NN algorithm by utilizing combined with the SAS-SV questionnaire. We expect that the system can help determine the extent of the influence of gadget addiction within four addiction levels: mild, normal, moderate and severe. SAS-SV is a short version of Smartphone Addiction Scale, an instrument to measure the Smartphone addiction developed by Korean National Information Society Agency (NIA).[6-9] SAS-SV consists of 6 factors (i.e., interference with daily living, positive anticipation, withdrawal, cyber-oriented relationships, overuse, and tolerance) with a five-point Likert scale (1: "strongly disagree" and 5: "strongly agree") based on self-reporting. Next, to verify the concurrent validity of SAS-SV based on the previously published internet addiction scale and mobile phone addiction scale, Smartphone addiction proneness scale (SAPS) was added. The SAPS consists of 15 items scored on a four-point Likert scale (1: "not at all" and 4: "always").[10]

Data mining is needed to support this research. Data mining is a process that uses statistical techniques, mathematics, artificial intelligence, and machine learning to extract and identify useful information and related knowledge from large databases [11]. The term data mining has the essence as a discipline whose main goal is to find, explore, or mine knowledge from the data or information that we have. KNN is one alternative which can be used to extract valuable information in the data mining process. In KNN or machine learning principles, the dataset will be divided into two groups: training-set and testing-set [ANA].[12] The training-set is a group of data used to teach (train) the machine learning algorithm, how to extract useful features, and finally to make a prediction model [Ana]. Whereas, the testing-set is a secondary (or tertiary) data group used to assess (test) the prediction model after it has been trained.[13]

The system development method used in this study is a prototype, a physical model of the system's work and serves as an initial version of the system [Ogedebe 2012]. In order for this prototyping process to work well, it is necessary to define the rules at an early stage: developers and users must have one understanding that the prototype was built to define initial requirements. The prototype will be removed or added to its part so that it is in accordance with the planning and analysis carried out by the developer until the trials are carried out simultaneously along with the development process [14]. We made the prototype of the system in the form of a website.

2. METHODS

The flowchart that describes the stage for making this prototype systems can be seen in **Figure 1**.

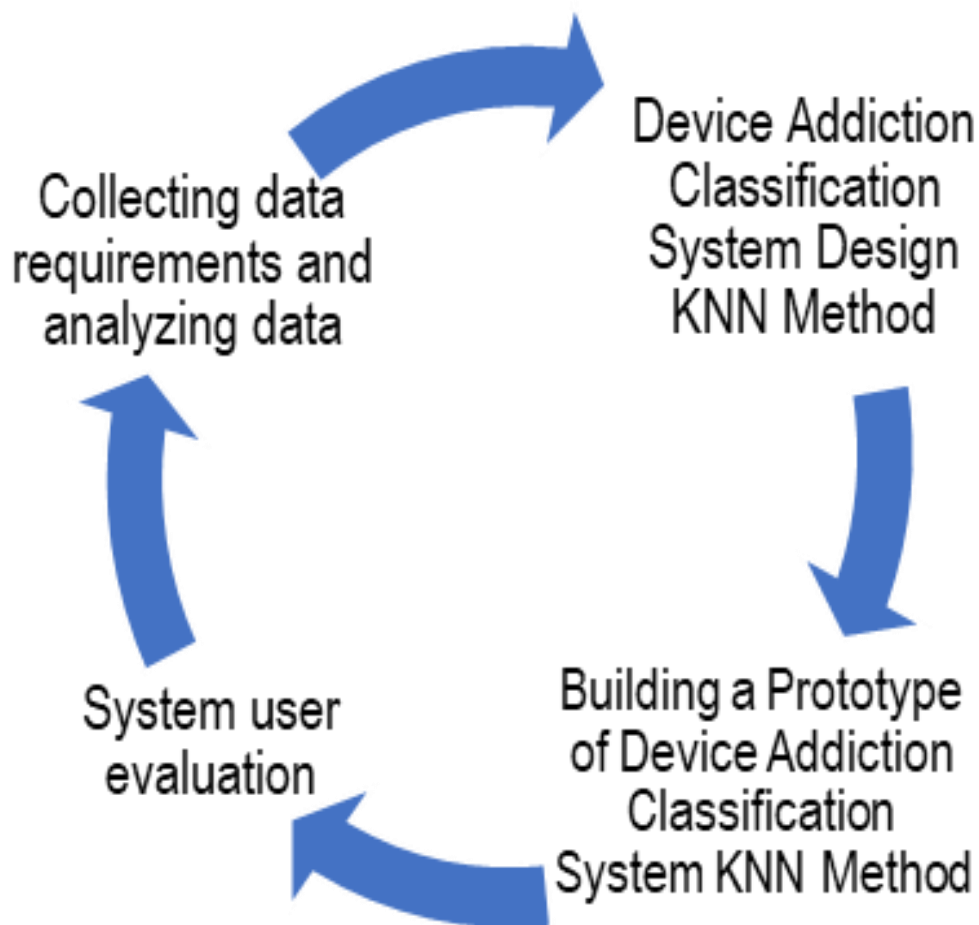


Figure 1. Research flowchart.

The following is an explanation of the above chart. Data collection here intends to meet the needs of the database of the program and the analysis phase of the researcher analyzes any data that has been collected whether it is in accordance with the needs or not. Such as preparing hardware that supports research and supporting software. At the stage of data collection, the research students collected 63 questionnaire data filled out by Informatics Engineering students and tested with a system of 15 data.

At the design stage, here the researcher begins to design the level of student addiction using KNN. For convenience, users can easily find out the level of addiction required by the application with the KNN algorithm.

The next stage is to build a prototype, user needs in more detail because users often have difficulty in delivering their needs in detail without seeing a clear picture. To anticipate that the project can run according to plan, at this stage the design that has been designed is then made a program that presents the classification system for gadget addiction using the KNN method.[15]

After the coding phase of the application is complete, this testing will be carried out to find errors from the application of the KNN algorithm application to measure whether the diagnosis produced by the expert is the same or not with the results given by the system. After completion, all stages of the research carried out have been completed.[16]

2.1. Use Case

Use case is a description of the actions performed by the system that produce a measurable result for an actor. The use case here consists of two, namely the Admin use case and the Student use case, which can be seen in **Figure. 2** and **Figure. 3**.

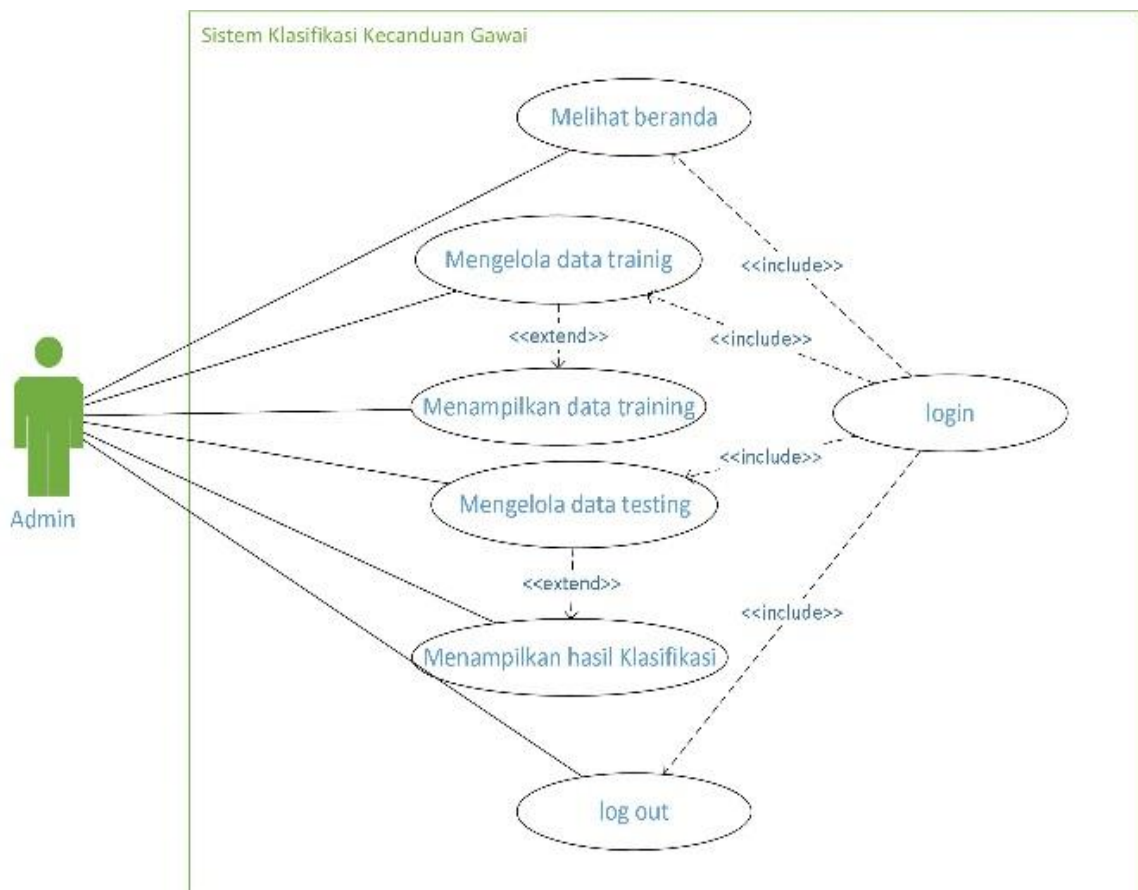


Figure 2. Admin use case diagram.

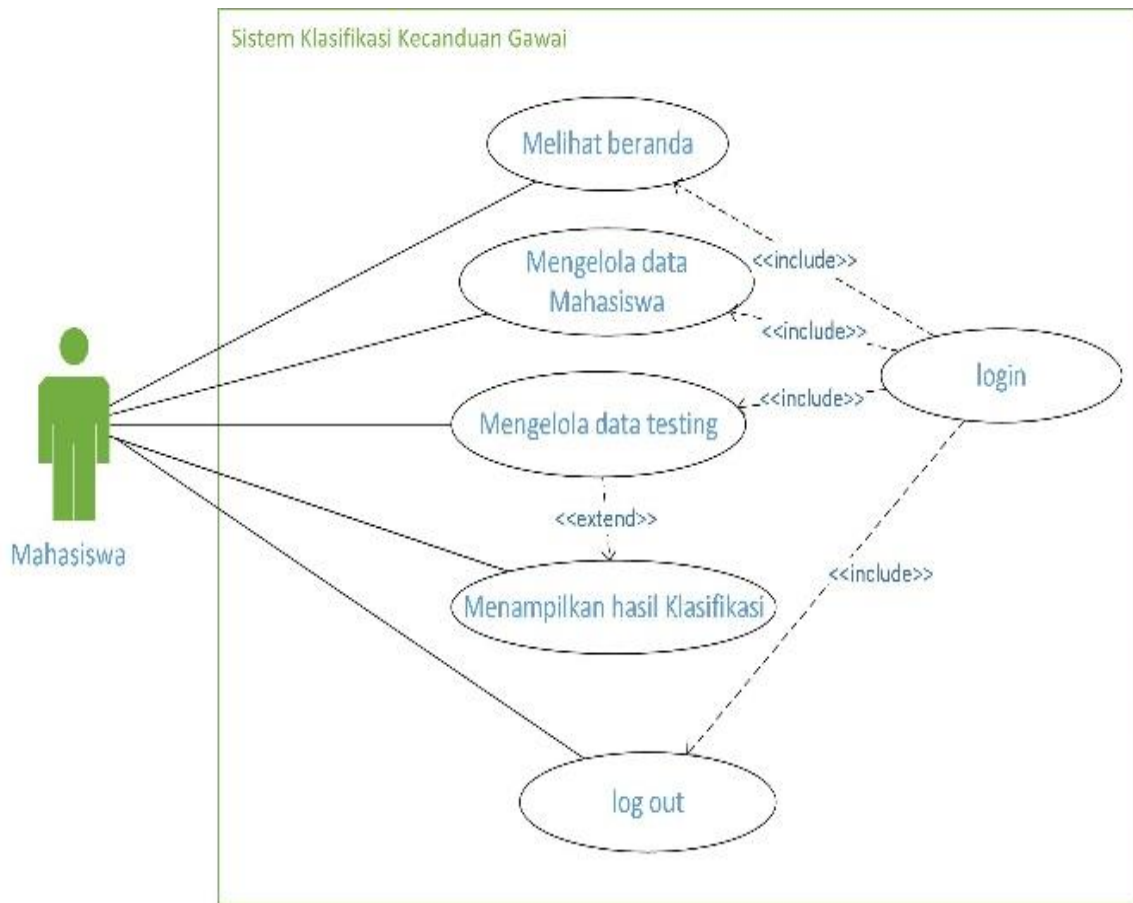


Figure 3. Student use case diagram.

2.2. K-Nearest Neighbor Method

There are several stages that will be carried out in the K-NN method starting with data collection, determining the value of the K parameter, calculating the distance using the Manhattan Euclidean formula.[17] How to calculate the K-Nearest Neighbor algorithm by using the distance calculation formulation The Manhattan distance Equation[1] is as follows:

$$D(x, y) = \sum_i^n |x_i - y_i| \tag{1}$$

Where n = large amount of data, i = represents the attributes used, xi = training data, yi = testing data.[18] The sample data is used as training data, the criteria used in classifying gadget addiction are five criteria can be seen in **Table 1**.

Table 1. Symptom Testing Data.

No	Symptom	Score
x ₁	I'm leaving my job because I'm using my smartphone for other things	3
x ₂	I feel calm or comfortable when using my device	2
x ₃	I feel excited when I use my device	1
x ₄	I feel confident when using my device	2
x ₅	I'm able to relieve stress by using my smartphone	1

Table 2. shows examples of training data from gadget addiction data. The data displayed is an example of training data and testing data on 4 tests.

Table 2. Example of Training Data.

Data	x ₁	x ₂	x ₃	x ₄	x ₅	Status
1	1	5	4	4	6	Low
2	3	2	3	5	6	Low
3	4	5	5	6	1	Medium
4	2	3	6	1	5	High
y	3	2	1	2	1	-

Determine the parameter K, here the value of K = 3 and calculate the square of the Manhattan distance of each object to the sample data given for the 1st data to the 4th data (data testing). Calculate new data distance with all training data (Manhattan distance)[19]:

$$\begin{aligned}
 \text{Data one using K-NN Method (1)} &= \sum_i^n |x_i - y_i| \\
 &= |x_{1_1} - y_1| + |x_{2_1} - y_2| + |x_{3_1} - y_3| + |x_{4_1} - y_4| + \\
 &\quad |x_{5_1} - y_5| \\
 &= |1 - 3| + |5 - 2| + |4 - 1| + |4 - 2| + |6 - 1| \\
 &= 2 + 3 + 3 + 2 + 5 \\
 &= 15
 \end{aligned}$$

$$\begin{aligned}
 \text{Data two using K-NN Method (1)} &= \sum_i^n |x_i - y_i| \\
 &= |x_{1_2} - y_1| + |x_{2_2} - y_2| + |x_{3_2} - y_3| + |x_{4_2} - y_4| \\
 &\quad + |x_{5_2} - y_5| \\
 &= |3 - 3| + |2 - 2| + |3 - 1| + |5 - 2| + |6 - 1|
 \end{aligned}$$

$$= 0 + 0 + 2 + 3 + 5$$

$$= 10$$

$$\begin{aligned} \text{Data three using K-NN Method (1)} &= \sum_i^n |x_i - y_i| \\ &= |x_{1_3} - y_1| + |x_{2_3} - y_2| + |x_{3_3} - y_3| + |x_{4_3} - y_4| \\ &\quad + |x_{5_3} - y_5| \\ &= |4 - 3| + |5 - 2| + |5 - 1| + |6 - 2| + |1 - 1| \\ &= 1 + 3 + 4 + 4 + 0 \\ &= 12 \end{aligned}$$

$$\begin{aligned} \text{Data four using K-NN Method (1)} &= \sum_i^n |x_i - y_i| \\ &= |x_{1_4} - y_1| + |x_{2_4} - y_2| + |x_{3_4} - y_3| + |x_{4_4} - y_4| \\ &\quad + |x_{5_4} - y_5| \\ &= |2 - 3| + |3 - 2| + |6 - 1| + |1 - 2| + |5 - 1| \\ &= 1 + 1 + 5 + 1 + 4 \\ &= 12 \end{aligned}$$

Order from lowest score to highest is data two with 10 point (low), data three with 12 point (medium), data four with 12 point (high), and data one with 15 point (low). Because $K = 3$, the 3 closest distance data taken as classification are data two with a value of 10 and a low classification, data three with a value of 12 with a medium classification and data four with a high classification.[20] From the sample 3 data, it is known that the result is low addition.

3. RESULTS AND DISCUSSION

3.1. Data Collection Process

In the process of collecting data, the researcher uses a coefficient to collect data. The researcher made a questionnaire using a Google form, then distributed it through The Unkhair Informatics Engineering internal group. From the results of the questionnaire distribution, the researcher got 78 questionnaire data that had been filled in, after the questionnaire data was filled in then the data was given to an expert on behalf of Amalia S.J Kahar, S. Psi., M. Psi., Psychologist to be diagnosed and used as training data by researchers. The results of the data collection that has been collected can be seen in **Table 3**.

Table 3. Total Data.

No	Statement	SS	S	AS	TS	STS
1	Due to the use of my smartphone, it is difficult to do work according to the schedule that I have set beforehand	6	10	16	20	11
2	I find it difficult to concentrate while in class, doing assignments, or working due to smartphone use	2	12	13	24	12
3	I feel pain in my wrists or back of neck when using my smartphone	5	16	11	23	8
4	I can't afford it if I am required to not have a smartphone	9	13	15	8	18
5	I feel impatient and restless when I'm not holding my smartphone	3	14	18	16	12
6	I think about my smartphone even when I'm not using it	1	12	14	22	14
7	I will never stop using my smartphone even though I know that my daily life has been greatly affected by smartphones	3	15	9	23	13
8	I check my smartphone regularly so I don't miss other people's conversations on social media	9	5	19	16	14
9	I always use my smartphone longer than I planned	15	19	16	9	5
10	People around me tell me that I use my smartphone excessively	8	7	15	20	13

3.2. K-Nearest Neighbors (KNN) Testing

In this test used 78 training data where mild addiction 31 moderate addiction 35 severe addiction 12 and 15 Testing data, the test is carried out using a questionnaire consisting of 10 questions where each question we can choose strongly agree, agree, somewhat agree, disagree, and strongly disagree. After calculating the 13 correct data and 2 incorrect data, using the value of $K = 3$ while the value of $K = 5$, 12 correct data and 3 incorrect data can be seen in **Table 4**.

Table 4. Comparison of Data Status Before and After Using the KNN Method.

No	Previous Status	Status After KKN with K = 3	Status After KKN with K = 5
1	Low Addiction	Low Addiction	Low Addiction
2	Low Addiction	Low Addiction	Low Addiction
3	Low Addiction	Low Addiction	Low Addiction
4	Low Addiction	Medium Addiction	Medium Addiction
5	Medium Addiction	Medium Addiction	Medium Addiction
6	High Addiction	High Addiction	High Addiction
7	Medium Addiction	Medium Addiction	Medium Addiction
8	High Addiction	High Addiction	Medium Addiction
9	Medium Addiction	Medium Addiction	Medium Addiction
10	Medium Addiction	Medium Addiction	Medium Addiction
11	Medium Addiction	Low Addiction	Low Addiction
12	Medium Addiction	Medium Addiction	Medium Addiction
13	High Addiction	High Addiction	High Addiction
14	High Addiction	High Addiction	High Addiction
15	Medium Addiction	Medium Addiction	Medium Addiction

3.3. System Interface

The User homepage display is a view that can be opened by users who are level as users who can only consult but cannot manage the data in it. The view of the user's homepage can be seen in **Figure. 4**.

**Figure. 4.** User Home Display.

The display of the consultation results is the result of the user conducting a consultation where the results show whether the user is addicted to gadgets. The display of the results of the consultation can be seen in **Figure. 5**.



Figure. 5. Consultation Results Display.

3.4. Analysis Results

In the research method, the results obtained from the beginning of making the system starting from collecting data requirements and analyzing data then after that the design of the gadget addiction classification system, the stages began and the design drawings were described using entity relationship diagrams, use case diagrams, activity diagrams, sequence diagrams, and flowcharts, where each diagram discusses the entirety of the system, how the actors operate the system that was built, an overview of the activities that occur in the system, what users do in the system, up to a description of what tables are displayed. is in the system database, how the tables can be related and an overview of the flow of the system being built. After the system is finished, the system testing process will be carried out, namely black box testing so that it is known whether the functions or features of the system are in accordance with what is desired or not. In testing the application of the KNN method, it shows that the KNN algorithm can be applied to measure the level of addiction to students in the Informatics Engineering Study program, Faculty of Engineering Unkhair with the results of testing the accuracy value for each k value with 78 training data and 15 testing data. different, namely the value of $K = 3$ obtained 86% accuracy. And the value of $k = 5$ obtained 80% accuracy.

4. CONCLUSION

Based on the discussion that has been described in chapter IV, the researchers can conclude that the research method results obtained from the beginning of making the system starting from collecting data requirements and analyzing data then after that the design of the gadget addiction classification system with the KNN method, shows that the KNN algorithm can applied to measure the level of addiction to students in the Informatics Engineering Study program, Faculty of Engineering Unkhair. With the results of testing the accuracy of each value of k with 78 training data and 15 test data (testing) it was found that for each different accuracy, the value of $K = 3$ obtained 86% accuracy. And the value of $K = 5$ obtained 80% accuracy.

In collecting data that has been managed by experts, it was found that 31 students experienced mild addiction, 35 students experienced moderate addiction and 12 students experienced severe addiction. Black box testing runs as expected by the system, in the form of testing the login process, registration, inputting training data and inputting testing data.

Then the researchers gave suggestions for students to change the pattern of excessive use of gadgets by simply using smartphones as needed according to their basic needs, namely for communication and information needs as necessary. The most important pattern to change is the duration of gadget use. Researchers also provide advice for health agencies to consider that technology is currently developing rapidly and its use can have an impact on health such as eyes, hearing, sleep patterns, and so on so that prevention is needed for this which can begin with impact assessment and then continue with counseling. The next researcher is expected to develop a system using different methods or combining the K-NN method with other methods in order to help improve accuracy results.

5. AUTHORS' NOTE

The authors declare that there is no conflict of interest regarding the publication of this article. Authors confirmed that the paper was free of plagiarism.

6. REFERENCES

- [1] Oktavia, N., dan Mulabbiyah, M. (2019). Gawai dan kompetensi sikap sosial siswa MI (studi kasus pada siswa kelas V MIN 2 kota Mataram). *El Midad*, 11(1), 19-40.
- [2] Abdi, S., dan Karneli, Y. (2020). Kecanduan game online: penanganannya dalam konseling individual. *Guidance*, 17(2), 9-20.
- [3] Widya, R. (2020). Dampak negatif kecanduan gadget terhadap perilaku anak usia dini dan penanganannya di paud Ummul Habibah. *Jurnal Abdi Ilmu*, 13(1), 29-34.
- [4] Sahambangung, J. R., Munir, R., dan Sanger, J. B. (2018). Sistem klasifikasi pengguna narkoba menggunakan metode k-nearest neighbor. *Jurnal Ilmiah Realtech*, 14(2), 169-179.
- [5] Sisbintari, K. D., dan Setiawati, F. A. (2021). Digital parenting sebagai upaya mencegah kecanduan gadget pada anak usia dini saat pandemi COVID-19. *Jurnal Obsesi: Jurnal Pendidikan Anak Usia Dini*, 6(3), 1562-1575.
- [6] Hasanah, U., Hijrianti, U. R., dan Iswinarti, I. (2020). Pengaruh smartphone addiction terhadap perilaku agresif pada remaja. *Proyeksi: Jurnal Psikologi*, 15(2), 182-191.
- [7] Putra, A. D. A., dan Juanita, S. (2021). Analisis sentimen pada ulasan pengguna aplikasi bibit dan bareksa dengan algoritma KNN. *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, 8(2), 636-646.
- [8] Naufal, M. F., Kusuma, S. F., Tanus, K. C., Sukiwun, R. V., Kristiano, J., dan Lieyanto, J. O. (2021). Analisis perbandingan algoritma klasifikasi citra chest x-ray untuk deteksi COVID-19. *Teknika*, 10(2), 96-103.
- [9] Angreni, I. A., Adisasmata, S. A., Ramli, M. I., dan Hamid, S. (2018). Pengaruh nilai k pada metode k-nearest neighbor (KNN) terhadap tingkat akurasi identifikasi kerusakan jalan. *Rekayasa Sipil*, 7(2), 63-70.
- [10] Arthy, C. C., Effendy, E., Amin, M. M., Loebis, B., Camellia, V., and Husada, M. S. (2019). Indonesian version of addiction rating scale of smartphone usage adapted from smartphone addiction scale-short version (SAS-SV) in junior high school. *Open Access Macedonian Journal of Medical Sciences*, 7(19), 3235-3239.

- [11] Nikmatun, I. A., dan Waspada, I. (2019). Implementasi data mining untuk klasifikasi masa studi mahasiswa menggunakan algoritma k-nearest neighbor. *Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer*, 10(2), 421-432.
- [12] Iskandar, J. W., dan Nataliani, Y. (2021). Perbandingan naïve bayes, SVM, dan K-NN untuk analisis sentimen gadget berbasis aspek. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 5(6), 1120-1126.
- [13] Naufal, M. F. (2021). Analisis perbandingan algoritma SVM, KNN, dan CNN untuk klasifikasi citra cuaca. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 8(2), 311-317.
- [14] Purnomo, D. (2017). Model prototyping pada pengembangan sistem informasi. *JIMP (Jurnal Informatika Merdeka Pasuruan)*, 2(2), 54-61.
- [15] Putra, A. D. A., dan Juanita, S. (2021). Analisis sentimen pada ulasan pengguna aplikasi bibit dan bareksa dengan algoritma KNN. *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, 8(2), 636-646.
- [16] Zaky, U., Prasetyoaji, A., dan Fathullah, I. (2020). Klasifikasi menggunakan metode naïve bayes: tingkat pengaruh penggunaan gadget terhadap kematangan kecerdasan emosi mahasiswa. *JUSIFO (Jurnal Sistem Informasi)*, 6(2), 101-110.
- [17] Halim, A. A. D., dan Anraeni, S. (2021). Analisis klasifikasi dataset citra penyakit pneumonia menggunakan metode k-nearest neighbor (KNN). *Indonesian Journal of Data and Science*, 2(1), 01-12.
- [18] Nasution, M. R. A., dan Hayaty, M. (2019). Perbandingan akurasi dan waktu proses algoritma K-NN dan SVM dalam analisis sentimen twitter. *Jurnal Informatika*, 6(2), 226-235.
- [19] Sahar, S. (2020). Analisis perbandingan metode k-nearest neighbor dan naïve bayes classifier pada dataset penderita penyakit jantung. *Indonesian Journal of Data and Science*, 1(3), 79-86.
- [20] Baita, A., Pristyanto, Y., dan Cahyono, N. (2021). Analisis sentimen mengenai vaksin sinovac menggunakan algoritma support vector machine (SVM) dan k-nearest neighbor (KNN). *Information System Journal*, 4(2), 42-46.