Collabits Journal
Vol. 1 No. 3 September 2024 : 186-192
E-ISSN : 3046-6709, P-ISSN : 3062-8601
https://publikasi.mercubuana.ac.id/index.php/collabits

Implementation of The Naive Bayes Algorithm on Online Game Addiction and Its Impact on Students

Thanel Richard Angwarmasse[1], Shafrizal Fadillah Akbar[2*]

Informatics Engineering Study Program, Mercu Buana University

*Coressponden Author: rizalakbar796@gmail.com

Abstract - This research aims to implement the Naive Bayes algorithm in analyzing online game addiction and its impact on individuals. Online gaming addiction has become a global phenomenon with significant psychological, social, and academic implications. For this reason, an effective analytical tool is needed to identify the factors that contribute to this addiction and its impact. The Naive Bayes algorithm was chosen because of its ability to carry out classification based on probability, which is very suitable for handling complex and diverse data. This research collects data from questionnaires that cover demographic aspects, frequency of play, duration of play, and perceived impact. The analysis results show that the Naive Bayes algorithm has quite high accuracy in classifying individuals who are addicted to online games. In addition, this study identified several key factors that are closely related to addiction, such as age, gender, and motivation to play. The most prominent impacts of this addiction include decreased academic performance, disturbed sleep patterns, and problems with social relationships. With the implementation of the Naive Bayes algorithm, it is hoped that it can contribute to prevention and early intervention efforts against online game addiction. This research also opens up opportunities for further development in the use of other machine-learning techniques for digital behavior analysis.

Keywords :
Naive Bayes;
Online Game Addiction;
Impact of Addiction;
Machine Learning;
Behavior Analysis;

Article History:
Received: 09-06-2024
Revised: 15-07-2024
Accepted: 28-08-2024

Article DOI : 10.22441/collabits.v1i3.27301

1. Introduction

The development of this increasingly advanced era is of course influenced by advances in information and communication technology, the use of which can now be felt and reached by various age groups. Mobile phones or gadgets, which initially could only be managed by adults, are now known and used by young children. This phenomenon occurs due to several factors, such as the busyness of parents which makes them often give gadgets to their children as a means of diversion, as well as the increasingly affordable prices of gadgets. Apart from that, the development of applications and content specifically for children that are educational and interesting also encourages the use of gadgets at an early age (Damayanti, 2023).

The internet plays an important role in our daily lives, both for important needs and simple needs. In the current era, the internet has become more advanced with rapid technological developments, so that more and more people use it for various purposes, including looking for entertainment. The internet not only makes access to information easier, but also provides a platform for faster and more efficient communication through instant messaging applications and social media (Hermawan & Kudus, 2021).

There are various forms of entertainment available via the internet and technological media, such as watching films via streaming services, playing on social media that allows interaction with friends and family, and listening to music from various digital platforms. One very popular form of entertainment via the internet is online games. Online games offer a variety of gaming experiences that can be easily accessed by anyone, including children who have been accustomed to using gadgets from an early age.

Online games not only provide entertainment, but also often contain educational elements that

can help children's cognitive development. However, on the other hand, there are challenges and risks that need to be taken into account, such as the potential for addiction, exposure to inappromente content, and lack of physical activity. Therefore, it is important for parents to monitor and regulate their children's use of gadgets and the internet, ensuring that technology is used in a healthy and productive way (Juliansyah, 2020).

Overall, advances in information and communication technology have changed various aspects of our lives, from the way we work, study, to the way we seek entertainment. It is important to continue to follow these developments wisely, maximizing the benefits while minimizing the potential risks.

2. Research Methodology

2.1 Research Type

This research uses a qualitative descriptive research approach, which utilizes qualitative data and describes it descriptively. Qualitative descriptive research is usually used to analyze events, phenomena, or social situations by displaying research results directly without changing the data (Yuliani, 2018). Apart from that, this research uses the Naive Bayes method. It is hoped that by using this method, this research can predict how much negative impact online games have on millennial teenagers and make it easier to classify them. Naive Bayes relies on the simplifying assumption that given output values, attribute values are conditionally independent of each other. In other words, the joint observing probability is the product of the individual probabilities (Watratan, Puspita B., & Moeis, 2020).

The advantage of the Naive Bayes method is that it only requires a small amount of training data to determine the parameter estimates required in the classification process. Training data is used as a reference for the calculations of each algorithm, while testing data is used to assess whether the predictions and determinations made by each algorithm are correct or not (Hidayat, Huda, & Pradita, 2024).

2.2 Types of Online Games

There are several types of online games, here are some examples :

1. Battle Royale
2. Multiplayar Online Battle Arena (MOBA).
3. Open world role-playing game (RPG)

2.3 Data collecting technique

During this research, the author generally used the following methods to collect data:

1) Observation: This observation method involves the author observing or reviewing online game players in the environment where the author lives.
2) Literature Study: The purpose of literature study is to obtain supporting data to support an argument. The author collected data by reading books, reference journals, filling out online questionnaires, and searching for articles on the internet.
3) Questionnaire: To obtain data, the author distributed a questionnaire in the form of questions about online games to respondents. The questionnaire, which the author created through the use of Google Form, is a data collection method that involves communication with the data source.

2.3 Data Analysis Techniques

Public perception of the rapid development of online games, which are increasingly being used, especially by young students, university students and the general public, is measured through data analysis. After the data is collected and processed, the Naive Bayes Algorithm is used to generate relevant comparison data.

3. Results and Discussion

Next, the naive Bayes algorithm will be used to process and analyze data that has been collected through online questionnaires distributed via social media. The data is divided into two criteria: training data and testing data. These two criteria are used to apply manual calculations, producing new information about whether online game users are still normal or addicted. Using Microsoft Excel, you can search for odds and calculate the odds for each attribute. Furthermore, to test the level of accuracy, Rapid Miner is used as a tool in the process of testing the accuracy level of the classification (Asmiati &

Fatmawati, 2020).

The following is a description of the accumulated data, the variables used to classify the data are as follows:

1. Full Name
   Variable containing the name of the online game user
2. Gender
   Variable to group data into 2 categories, Men and Women

Tabel I
GENDER

Attribute	Criteria	Amount	Classification	
Gender	Men	33	Addicted	13
			Normal	20
	Women	7	Addicted	3
			Normal	4

3. Age
   Variables to group data into 2 categories, from age 12-17 years old and age 18-22 years old

TABEL II
AGE

Attribute	Criteria	Amount	Classification	
Age	12-17 Years Old	6	Addicted	0
			Normal	6
	18-22 Years Old	34	Addicted	16
			Normal	18

4. Amount of Days
   Variables to group data into 4 categories, namely 1-2 Days, 3-4 Days, 5-6 Days, and Every Day.

TABEL III
AMOUNT OF DAYS

Attribute	Criteria	Amount	Classification	
Amount of Days	1-2 Days	7	Addicted	1
			Normal	6
	3-4 Days	5	Addicted	3
			Normal	2
	5-6 Days	3	Addicted	1
			Normal	2
	Setiap Days	25	Addicted	11
			Normal	14

5. Durasi waktu
   Variables to group data into 3 categories, 1-2 Hours, 3-4 Hours, and 4+ Hours.

TABEL IV
DURASI WAKTU

Attribute	Criteria	Amount	Classification	
Duration	1-2 Hours	14	Addicted	5
			Normal	9
	3-4 Hours	11	Addicted	3
			Normal	8
	4+ Hours	15	Addicted	8
			Normal	7

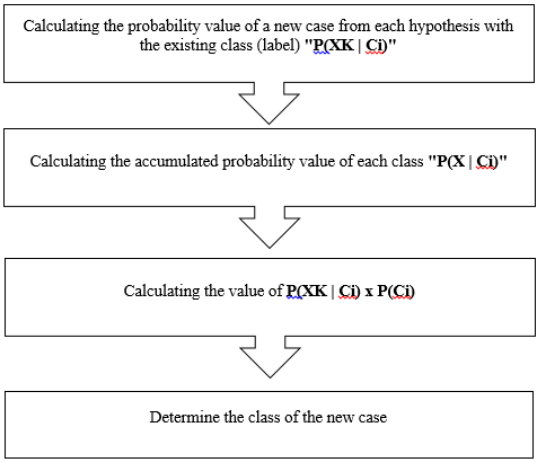
6. Classification
   Class attribute consisting of Addicted and Normal variables

TABEL V
CLASSIFICATION

Attribute	Criteria	Amount
Classification	Addicted	16
	Normal	24

3.1 Initial Data Processing

Following are the steps for solving the problem using the Naïve Bayes algorithm:



Fiigure 1 Naive Bayes Algorithm Flow

- Calculate the values of “P(XK|Ci)”

P(Criteria 1 = Men | Class. = “Addicted”)
P(Criteria 1 = 13/33) = 0.393
P(Criteria 1 = Men | Class. = “Normal”)
P(Criteria 1 = 20/33) = 0.607

P(Criteria 2 = 12-17yo | Class. = “Addicted”)
P(Criteria 2 = 0/6) = 0
P(Criteria 2 = 12-17yo | Class. = “Normal”)
P(Criteria 2 = 6/6) = 1

P(Criteria 3 = Every Day | Class. = “Addicted”)
P(Criteria 3 = 11/25) = 0.44
P(Criteria 3 = Every Day | Class. = “Normal”)
P(Criteria 3 = 14/25) = 0.56

P(Criteria 4 = 4+ Hours | Class. = “Addicted”)
P(Criteria 4 = 8/15) = 0.533
P(Criteria 4 = 4+ Hours | Class. = “Normal”)
P(Criteria 4 = 7/15) = 0.467

- Calculate the accumulated opportunity value for each class of "P(X|Ci)"

P(X | Class. = “Addicted”)
= 0.393 x 0 x 0.44 x 0.533
= 0
P(X | Class. = “Normal”)
= 0.607 x 1 x 0.56 x 0.467
= 0.15874264

- Calculate the value of P(X|Ci) x P(Ci)

P(X|Class=“Addicted”) x P(Class=“Addicted”)
= 0 x 16/40
= 0 x 0.4
= 0
P(X|Class=“Normal”) x P(Class=“Normal”)
= 0 x 24/40
= 0.15874264 x 0.6
= 0.095245584

- Determine the class of the new case

Based on the calculations that have been carried out, it can be seen that the value of P(X|Class="Normal") x P(Class="Normal") is higher than the value of P(X|Class="Addicted") x P(Class="Addicted ”) with a comparison value = **0.095245584** versus **0**, while the classification result of the addicted class is **0.4** while the normal class has a value of **0.6**. Therefore, it can be concluded that the surveyed online game users do not experience addiction when playing online games, or we can state it as **Normal**.

Data Processing Using MS-Excel

First of all, we have to prepare the dataset needed to implement the Naïve Bayes algorithm. Then start calculating the probability values for each class, here is the formula:

$$P(h|x) = \frac{P(x|h) \times P(h)}{P(x)}$$

To apply the formula above, you need the COUNTIF function, COUNTA function, and mounting function in Microsoft Excel.

= COUNTIF(Range of classification attribute column, "Search criteria")

Example:

= COUNTIF(G2:G41, “Satisfied”)

= COUNTA(Range of classification attribute columns)

= SUM(Column Cells to be calculated)

Here are the results:

- ✓ Addicted = 16 People
- ✓ Normal = 24 People
- ✓ Total data = 40 People

We convert the data above into percentage form with the formula:

$$\frac{\text{Total Attribute}}{\text{Total data}} \times 100\%$$

So the results are:

TABEL VI

PROBABILITAS CLASS

Probabilitas Class	
Classification	Nilai
Addicted	40%
Normal	60%
Total	100%

Next, we calculate the probability value of each attribute in the data set that we have. There are 4 data attributes, namely: Gender, Age, Amount of Days, Duration. The following is the formula for calculating the probability of each attribute:

$$P(x) = \sum_{n-1}^{n} P(x|hj)\, P(h)$$

To apply the formula above, we use the COUNTIFS function. The following is the formula for implementing the function:

= COUNTIFS(Range of the First attribute classification column, "The first criteria to be searched for from that range", Range of the Second attribute classification column, "The second criteria to be searched for from that range")

Example:

= COUNTIFS(G2:G41, "Unsatisfied", C2:C41, "Men")

After the amount of data is obtained, we convert it to percentage form using the formula that was previously used. The following are the results of calculating the probability of each attribute in tabular form:

TABEL VII
GENDER

Gender	Addicted	Normal
Men	81.25%	84%
Women	18.75%	16%
Total	100%	100%

TABEL VIII
AMOUNT OF DAYS

Amount hari	Addicted	Normal
1-2 Days	6.25%	25%
3-4 Days	18.75%	8.33%
5-6 Days	6.25%	8.33%
Everyday	68.75%	58.34%
Total	100%	100%

TABEL X
DURATION

Duration	Addicted	Normal
1-2 Hours	31.25%	37.5%
3-4 Hours	18.75%	33.34%
4+ Hours	50%	29.16%
Total	100%	100%

3.2 Implementation using Rapid Miner

The following is an implementation of the Naïve Bayes algorithm using Rapid Miner.

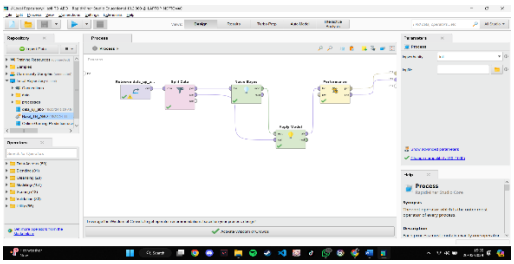


Figure 2 Naive Bayes Algorithm Design

In Figure 2 above is the design form of the Naïve Bayes algorithm model using 40 data that has been collected and then applied to the Rapid Miner tool.

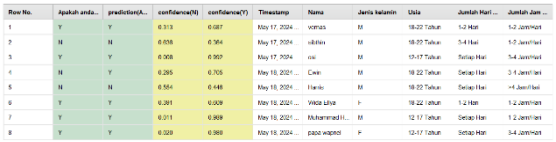


Figure 3 Test Data Result

In Figure 3 above, the results of the test data are shown with an amount of 8 because there are 40 data so it is divided into two, namely 32 for test data and 8 for test data..

3.3 Rapid Miner Test Results

The following are the results of tests that have been carried out using the Rapid Miner tool:

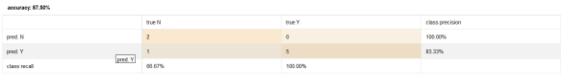


Figure 4 Performance Vector

The overall accuracy result from 40 data is 87.50%
Class Recall N 66.67%
Class Recall Y 100.00%
Class Precision N 100.00%
Class Precision Y 83.33%

PerformanceVector

```
PerformanceVector:
accuracy: 87.50%
ConfusionMatrix:
True:    N       Y
N:       2       0
Y:       1       5
```

Figure 5 Description Performance Vector

In Figure 5 it is explained that the vector performance accuracy is 87.50%

True:	N	Y
N :	2	0
Y:	1	5

ANALYSIS RESULTS

Based on the results of the research that has been carried out, new patterns, data and knowledge were obtained in the data processing process to classify the level of addicted use of online games. The aim of this data processing process is to produce data mining calculation patterns consisting of training and assessment data and to determine the possibility that each attribute will produce new information about whether many millennial teenagers are addicted to online games. Next, Rapid Miner is used to test the level of accuracy of the classification.

The results shown by the data mining calculation process using the manual method using the Naïve Bayes algorithm, show that the value of P(X | class. "Normal") x P(class. "Normal") is higher than P(X|class. "Addicted") x P(class. "Addicted"), namely = **0.095245584** versus **0**. Meanwhile, the results from the addicted class classification are worth **0.4** and the results from the normal class classification are worth **0.6**.

With an overall accuracy of 87.5%, a normal recall class of 66.67%, a normal recall class of 100%, and an addicted recall class of 66.67%, the research results of millennial teenagers who play online games are declared **NORMAL**, based on the results of data mining calculations and the accuracy level testing process using rapid miner.

4. SUMMARY

Based on the results of research and discussions regarding the classification of the Naive Bayes algorithm for online game users, it can be concluded that this method is easier for beginners to use. The results of data processing using the Rapid Miner application which uses the Naive Bayes method show that from 40 samples of online game user data, this algorithm is still categorized as Normal with an overall accuracy of 87.5%, Class recall addicted 66.67%, Class recall normal 100%, Class precision addicted 100% , and Class precision is normal 83.33%.

REFERENCES

Asmiati, N., & Fatmawati. (2020). Penerapan Algoritma Naive Bayes Untuk Mengclassification Pengaruh Negatif Game Online Bagi Remaja Milenial. *JTIM: Jurnal Teknologi Informasi dan Multimedia Vol. 2, No. 3*, 141-149.

Damayanti, D. (2023, Juli 13). *6 Alasan Orangtua Memberikan Gadget pada Anak Age Dini, Sibuk?* Retrieved from IDN TIMES JATENG: https://jateng.idntimes.com/life/family/daysdesy/6-alasan-orangtua-memberikan-gadget-pada-anak-age-dini-sibuk-c1c2?page=all

Hermawan, D., & Kudus, W. A. (2021). PERAN ORANG TUA DALAM MENCEGAH ANAK ADDICTED BERMAIN GAME ONLINE DI ERA DIGITAL. *Jurnal Pendidikan Indonesia Vol. 2 No. 5*, 778-789.

Hidayat, R., Huda, M., & Pradita, N. (2024). Comparison Of The C4.5 And Naïve Bayes Algorithm For Recommendations For Aid Recipients For The Smart Indonesian Program. *Kesatria: Jurnal Penerapan Sistem Informasi (Komputer dan Manajemen)*, 345-358.

Juliansyah, M. A. (2020). Makna dan Alasan Dibalik Orangtua Memberikan Fasilitas Gadget Kepada Anak Age Dini. *JPRMEDCOM: Journalism, Public Relation and Media Communication Studies Journal Vol. 2, No. 2*, 20-34.

Watratan, A. F., Puspita B., A., & Moeis, D. (2020). Implementasi Algoritma Naive Baiyes Untuk Memprediksi Tingkat Penyebaran Covid 19. *JACOST: Journal of Applied Computer Science and Technology Vol 1 No 1*, 7-14.

Yuliani, W. (2018). METODE PENELITIAN DESKRIPTIF KUALITATIF DALAM PERSPEKTIF BIMBINGAN DAN

KONSELING. *Quanta Journal Vol. 2 No. 2*, 83-91.