

---

## ACCURACY COMPARISON OF DATA MINING METHODS FOR INTERNET GAMING DISORDER CLASSIFICATION

**Juwita, Junidar, Viska Mutiawani, Zahnur Nurdin, Zul Ikram Zainuddin**

Faculty of Mathematics and Natural Science, Syiah Kuala University, Indonesia

Email: junidar678@unsyiah.ac.id

---

### Abstract

Excessive access to online games by players can lead to addiction. Based on the value of the answers in the questionnaire a person can be categorized into groups, mild addiction, moderate addiction and severe addiction. Generally, this assessment calculation is done manually or through procedural programming languages. This is not efficient for processing more and more data, for that the machine learning classification model is applied to solve the problem of program code repetition. This study compared the performance of three machine learning methods against two different types of questionnaires, namely questionnaires with Likert scale and questionnaires with yes no type. The case study used in this study is online game addiction among high school students in Banda Aceh City, Indonesia. This research successfully proved that the algorithm ... It is better to use questionnaires with data types...., while algorithms....are better to use for questionnaires with types.....with the accuracy of the three algorithms are as follows. This study reveals the emergence of online game addiction, especially among high school students within Banda Aceh city. The results depicted that as many as 6% of high school students in the city of Banda Aceh indicated experiencing online game addiction based on their reports. Another objective of this research is to find the best accuracy between Naive Bayes and Support vector machine (SVM) in classifying the severity of the online game. It found that SVM accuracy was higher than Naive Bayes for the case of online game addiction level classification in high school students in Banda Aceh. This study provided baseline data for further research.

**Keywords:** oline game addiction; support vector machine; naïve bayes

---

### INTRODUCTION

Internet gaming disorder (IGD) has been included as pathological behavior in the newest version of the Diagnostic and Statistical Manual of Mental Disorders (DSM) (5th ed. [DSM-5]) as been consented by American Psychiatric Association (APA) (Weiner & Craighead, 2010). A gaming disorder is a condition where players place gaming activity as a priority over other interests and daily exercises, and continuation of gaming even though it causes negative consequences (Aymé et al., 2015). Damage of gaming on human's mental health cannot be determined by longevity time playing, but rather depends on unexpected variables such as player characteristics, the diversion highlights, and the encounter of play . Gaming can become pathological when the playing becomes impairing, harming a person's social, job, family, education, and psychological functions (Gentile et al., 2011). Although digital gaming may drive a positive effect on well-being Johnson et al.,( 2013) some evidence has shown that online gaming is related to addictive behaviors (Chadwick & Wesson, 2016; Mandryk et al., 2020; Ng & Wiemer-Hastings, 2005; Van Rooij et al., 2011).

Globally, several instruments to measure the severity of internet gaming addiction have been carried out (Andreassen et al., 2012; Jiang & Jiang, 2019; Pontes & Griffiths, 2015). Despite their popularity, customization of such instruments is still necessary because understanding the cultural context of respondents is a concern (Jap et al., 2013). Have

succeeded to develop an online addiction questionnaire by adjusting some terms to become understandable by Indonesian students. Because one of the purposes of the current research was to investigate the internet gaming disorder among secondary school pupils in Banda Aceh, thus the Jap's questionnaire was used. Based on the total score for each answer, the addiction level consists of severe gaming addiction, mild gaming addiction, and normal gaming. Another goal set was to analyze the accuracy of the data mining methods in classifying three-level internet gaming addiction based on the data gathered. The popular classification techniques were Naive Bayes (NB) and Support Vector Machine (SVM). To date, there is no standard to choose which types of classifiers are most suitable to solve a problem, therefore it is necessary to test several types of classifiers to discover the best accuracy. Classifier accuracy is determined by several parameters including true positive rate (TP), false positive rate (FP), and precision (Narayan, 2021). This study will use these parameters in comparing the two classifiers, namely SVM and NB. This research provides valuable best practices for choosing a data mining classifier for a similar case. In addition, it also provides a summary of a statistic of internet gaming disorder among secondary school pupils within Banda Aceh city as initial data for similar research.

## **METHOD RESEARCH**

This research was a combination of exploratory case study research because it aims to find answers to the question "what" Yin, (2009) and experimental study to compare the performance of a data mining algorithm. This study aimed to explore the best suited algorithm for each internet gaming disorder classification based on respondents' self-answer. The samples of this research were students in secondary level in Banda Aceh City, Indonesia. Gender differences were not encountered in the classifier evaluation. Next Naïve Bayes and Support Vector Machine Algorithms were used to generate the model from a dataset. Then those models were analyzed for their accuracy in classifying the level of internet gaming disorder.

### **a. Questionnaires**

The questionnaire applied was constructed by (Xiao et al., 2011). The questions consisted of 7 items. It applied a 5-point Likert scale. It has adequate psychometric attributes for research use. Each question is presented in table 1

**Table 1. Questionnaire of Online Addiction for Indonesian Students**

Questions	Answer Scale				
	Never	Seldom	Sometimes	Often	Very
I think about playing an online game all					
My online game play time increases (for					
I play online game to run away from					
Others fail when they try to reduce my					
I feel uncomfortable if I cannot play					
Online game made my relationship with					
The time spent to play online games made					

b. Sampling Participants

Totally 359 respondents fulfilled the requirement for this research. The students who participated in this study had been playing online games minimally once within the past month and had not yet decided to stop (Xiao et al., 2011). Those students came from three public high schools located in Banda Aceh, the capital of Aceh Province, Indonesia. The age range of participants was Mehroof & Griffiths,( 2010) to Chiplunkar & Fukao, (2020) years old. Furthermore, the influence of other variables such as school location, age, and gender of the respondents was neglected.

c. Preprocessing

Before testing data to the classifier, that data was cleaned from the missing value and then it was divided into groups namely training data and testing data. Furthermore, the collected surveys then were calculated using clinical cut-off appraise; below 14 indicating normal gamer, scores of 14 to 21 indicating mild online game addiction, and score of 22 and above indicating strong online game addiction. After that, the questionnaire data were separated into 30% testing data and 70% training data. The data selection was conducted randomly.

d. Classification

In this study, the classification test was carried out using 2 data mining methods, namely Naive Bayesian and Support Vector Machine (SVM). Classification using Naive Bayesian and SVM methods was run using the WEKA application. The results to be compared from the stage are the confusion matrix, the precision value, the recall value, and the F-measure value. Evaluation is done to find out the accuracy results of the experiments which had been carried out. After the accuracy value is obtained, the validation process is conducted to get the best accuracy value using a confusion matrix and ROC curves. ROC curve visualized the positive and negative result divided by threshold value 25

## RESULT AND DISCUSSION

### Respondents' Profile

Researchers obtained 358 questionnaires that have been filled out by targetted samples. All samples are high school students with an age range of 15-17 years old. The proportion of sex in this study was 44% female and 56% male as shown in Figure 1 below:

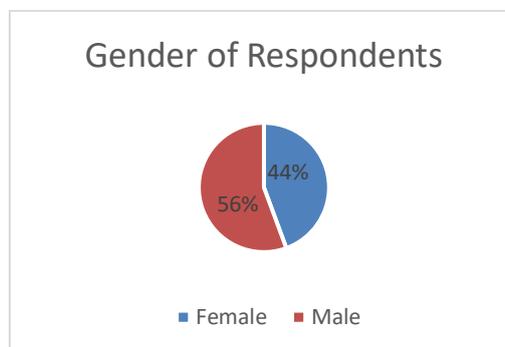


Figure 1. Gender of Respondents

The calculation results of the score show that 6.1 % of students are classified into severe addiction, 33% mild addiction, and 59% normal gamers. The pie chart is provided in Figure 2.

While the details of the severity of online game addiction according to gender are presented in Figure 3.

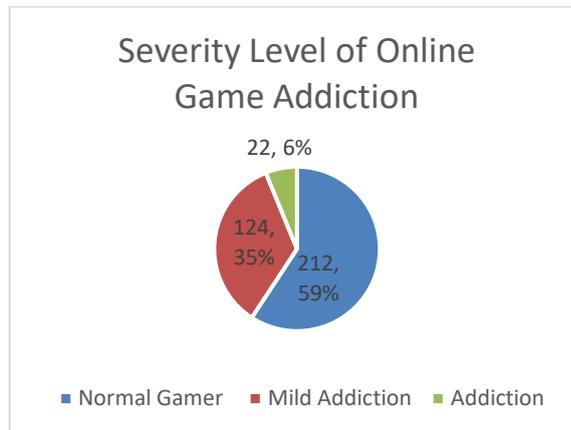


Figure 2. Severity Level of Online Game Addiction among Senior High School Students

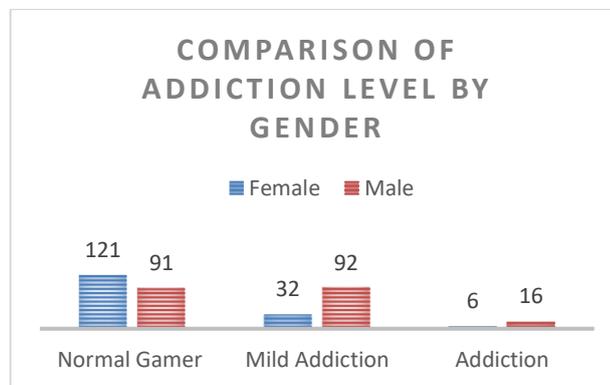


Figure 3. Addiction Level by Gender

Figure 3 shows that the number of male students who are slightly addicted to online games outnumbered female students. In general, 33.8% of female students who also play online games are still in the normal category. Although the proportion of female students addicted to online games is relatively low, specifically 1.4%, this already can be utilized as a baseline for further research in a similar area. On the other hand, the prediction of online game addiction cases (6%) among students in Banda Aceh were quite warning. This study may contribute by supporting baseline data regarding online game addiction for further diagnosis treatment.

### Classification

From 359 data collected, there were 9 missing values for the age column, as shown in Figure 4:

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 359 entries, 0 to 358
Data columns (total 9 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Umur                   350 non-null   float64
1   PreOccupation          358 non-null   float64
2   Tolerance               358 non-null   float64
3   Mood Modification      358 non-null   float64
4   Relapse                 358 non-null   object
5   Withdrawal              358 non-null   float64
6   Conflict                 358 non-null   float64
7   Sleep Deprivation      358 non-null   float64
8   Class                   358 non-null   object
dtypes: float64(7), object(2)
memory usage: 25.4+ KB
```

Figure 4. Dataset Details

The missing value is handled by entering the average age. Then the class division of training data and testing data were carried out. 70% of data was prepared for training while 30% others were prepared for testing. The details of the distribution are presented in Table 2.

**Table 2. Composition of training data and testing data**

		Number	Percentage (%)
Original Data	Normal Gamer	212	59
	Mild Addiction	124	35
	Addiction	22	6
Training Data	Normal Gamer	153	61
	Mild Addiction	87	35
	Addiction	11	4
Testing Data	Normal Gamer	59	55
	Mild Addiction	37	34
	Addiction	11	10

The results of the program output for 7 items online game Addiction questionnaire data classification using the Naïve Bayes method and using SVM, respectively, are shown in Tables 3 and 4 below.

**Table 3. Accuracy, F-Measure for Naïve Bayes**

Naïve Bayes Accuracy: 94 %

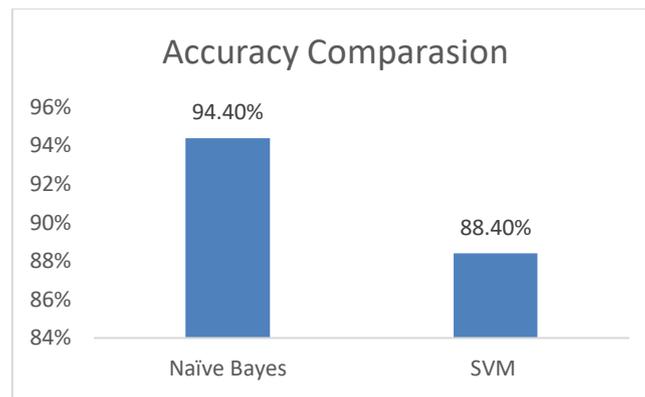
TP Rate	FP Rate	F-Measure	ROC Area	Class
0,95	0,00	0,89	1,00	Normal Gamer
1,00	0,08	0,78	0,99	Mild Addiction
0,72	0,00	0,31	0,99	Addiction

**Table 4. Accuracy, F-Measure For SVM**

SVM Accuracy: 88,89 %

TP Rate	FP Rate	F-Measure	ROC Area	Class
0,91	0,02	0,94	0,95	Normal Gamer
0,97	0,15	0,85	0,74	Mild Addiction
0,55	0,00	0,71	0,59	Addiction

From Tables 3 and 4 above, it can be seen that the Naive Bayes model divides positive and negative values. It is indicated by the ROC value equal to 1.0 for normal class gamers. Then the test results of the Naive Bayesian and SVM data mining classification methods on the 7-Item Online Game Addiction questionnaire were 94.4% and 88.8%, respectively. The comparison can be seen in figure 5 below:



**Figure 5. Comparison of Accuracy between Naïve bayes and SVM**

Based on the results of the accuracy comparison, it can be seen that the Naive Bayes classifier has better performance than SVM.

## CONCLUSION

The findings show that 6% of students in the city of Banda Aceh are indicated to be addicted to online games based on measurements through a questionnaire. Naive Bayes and SVM classifiers were used to classify the level of online game addiction. Naive Bayes show 94,4% accuracy which outperformed SVM with 88,40% accuracy.

## REFERENCES

- Andreassen, C. S., Griffiths, M. D., Hetland, J., & Pallesen, S. (2012). Development of a work addiction scale. *Scandinavian Journal of Psychology*, *53*(3), 265–272.
- Aymé, S., Bellet, B., & Rath, A. (2015). Rare diseases in ICD11: making rare diseases visible in health information systems through appropriate coding. *Orphanet Journal of Rare Diseases*, *10*(1), 1–14.
- Chadwick, D., & Wesson, C. (2016). Digital inclusion and disability. *Applied Cyberpsychology: Practical Applications of Cyberpsychological Theory and Research*, 1–23.
- Chiplunkar, N. N., & Fukao, T. (2020). *Advances in Artificial Intelligence and Data Engineering: Select Proceedings of AIDE 2019* (Vol. 1133). Springer Nature.
- Gentile, D. A., Choo, H., Liau, A., Sim, T., Li, D., Fung, D., & Khoo, A. (2011). Pathological video game use among youths: A two-year longitudinal study. *Pediatrics*, *127*(2), e319–e329.
- Jap, T., Tiatri, S., Jaya, E. S., & Suteja, M. S. (2013). The development of Indonesian online game addiction questionnaire. *PloS One*, *8*(4), e61098.
- Jiang, Q., & Jiang, Q. (2019). Risk Factors and Clinical Assessment of Internet-Addicted Adolescents. *Internet Addiction Among Cyberkids in China: Risk Factors and Intervention Strategies*, 85–142.
- Johnson, D., Wyeth, P., & Sweetser, P. (2013). The people-game-play model for understanding videogames' impact on wellbeing. *2013 IEEE International Games Innovation Conference (IGIC)*, 85–88.
- Mandryk, R. L., Frommel, J., Armstrong, A., & Johnson, D. (2020). How passion for playing World of Warcraft predicts in-game social capital, loneliness, and wellbeing. *Frontiers in Psychology*, *11*, 2165.
- Mehroof, M., & Griffiths, M. D. (2010). Online gaming addiction: the role of sensation seeking, self-control, neuroticism, aggression, state anxiety, and trait anxiety. *Journal of*

- Cyberpsychology, Behavior and Social Networking*, 13(3), 313–316.  
<https://doi.org/10.1089/CYBER.2009.0229>
- Narayan, Y. (2021). Comparative analysis of SVM and Naive Bayes classifier for the SEMG signal classification. *Materials Today: Proceedings*, 37, 3241–3245.
- Ng, B. D., & Wiemer-Hastings, P. (2005). Addiction to the internet and online gaming. *Cyberpsychology & Behavior*, 8(2), 110–113.
- Pontes, H. M., & Griffiths, M. D. (2015). Measuring DSM-5 internet gaming disorder: Development and validation of a short psychometric scale. *Computers in Human Behavior*, 45, 137–143.
- Van Rooij, A. J., Schoenmakers, T. M., Vermulst, A. A., Van Den Eijnden, R. J. J. M., & Van De Mheen, D. (2011). Online video game addiction: identification of addicted adolescent gamers. *Addiction*, 106(1), 205–212.
- Weiner, I. B., & Craighead, W. E. (2010). *The corsini encyclopedia of psychology, volume 4* (Vol. 4). John Wiley & Sons.
- Xiao, X., Wu, Z.-C., & Chou, K.-C. (2011). A multi-label classifier for predicting the subcellular localization of gram-negative bacterial proteins with both single and multiple sites. *PloS One*, 6(6), e20592.
- Yin, R. K. (2009). Case Study Research: Design and Methods, Essential guide to qualitative methods in organizational research. *Applied Social Research Methods Series*, 219(2), 212–259.

**Copyright holders:**

**Juwita, Junidar, Viska Mutiawani, Zahnur Nurdin, Zul Ikram Zainuddin (2023)**

**First publication right:**

**Injurity - Interdisciplinary Journal and Humanity**



**This article is licensed under a Creative Commons Attribution-ShareAlike 4.0 International**