The Impact of Video Games on University Students

Data Mining to Understand the Influence of Video Games on Academic Performance, Behavior, and Cognitive Development in Students

Walaa Hussain Ali Ahmed
202104424
Section 2

Fatema Hussain Habib
20192408
Section 2

Zahra Abdulghani Mansoor
20164923
Section 2

Manar Eyad Ahmed
202006306
Section 2

supervised by: Dr. Ahmed Zeki

***Abstract*—Video game addiction has a significant effect on students in universities. The purpose of this study is to use data mining to determine the association between video game addiction and the rate of students. Through this study, it will become clear to the student that addiction to video games affects his studies and GPA. Decision trees and Apriori are two of the data mining techniques employed in this study**.

INTRODUCTION

There are many behaviors that can affect a student's performance in several aspects, such as study, health, and rest. One of the most important reasons is addiction to video games, which can occupy a large area of the student's mind and negatively affect his university performance. Games prevent him from concentrating on lectures while explaining the doctor or following lessons. Also, this leads to students not finishing their Assignments and not being able to study well for exams. Therefore, data mining will help to discover the impact of video games on students' performance and GPA, and then determine whether video games affect students negatively or positively.

Data mining is the process of searching for information from large data and then analyzing it to discover patterns and find out useful information. [1]

DATA COLLECTION

Microsoft Forms was used as a platform to help collect data in order to learn more about video games and the performance of college students. We gathered 83 replies from university students who participated in the survey's questions.

DATA PREPARATION

The data was collected in an Excel file and then modifications were made to the attribute name to make it short. And the data was complete in the file without any Missing value. Also, most of the attribute value contains an answer (yes or no). After that, the file was converted to (.csv) to be used in the WEKA program.

SIMILAR WORK (KNOWLEDGE)

Essay on the effect of video games on academic performance [2]. It is one of the topics similar to our topic, where the effect of video games on student performance was studied. It was also studied whether these games affected students negatively or positively. In this article, one of the aspects that has been focused on is the effect of playing for long periods of time on academic assignments and GPA. On the other hand, others have supported the idea that those games have a positive impact on the development of students through solving puzzles and developing their language. However, the "participants" method was used to collect samples and then analyze them. Next, results showed that most of the students who had played video games had significantly lower GPAs than participants who indicated that they had not played video games.

In the second article, the effect of video games on academic performance and sleep duration in medical students [3]. It is also one of the topics similar to our topic in terms of education and health. The study also focused on evaluating the effect of regular video games on academic performance and sleep duration in healthy adolescents who indulge in video games. The GIMS method was used in this study, which was about the top students in the academy, about 50 students who play video games addictively and 50 who do not play video games. The results after collecting the data showed that video games have a significant negative impact on medical students, as they reduce the number of hours of sleep, which leads to fatigue and possible serious health problems.

ATTRIBUTES

The following nine attributes are part of the dataset created from the questionnaire given to university students: all of them are nominal. A brief description of each one, along with its values, is provided in the (Table 1 and Figure1) below.

No.		Name
1		Student_Age
2		Student_Gender
3		DailyGamingHours
4		TimeManagementSkills
5		GamingVSStudying
6		GamingAndFocus
7		GamingAndPriortization
8		GamingAndPhysicalHealth
9		GPA_Impact

Figure 1 Attributes in weka

Attributes	Description	Values
Student_Age	Identify students age	• 18-19 • 20-21 • 22-23 • 24-25
Student_Gender	Identify student gender	• Male • Female
DailyGamingHours	To determine how much hours each student spends playing per day	• 1-2 • 3-4 • 5-6 • 7-8 • Above 10
TimeManagementSkills	To determine whether the student can balance his time between studying and playing	• Yes • No
GamingVSStudying	To determine if the student has ever played games instead of studying for upcoming exams or completing assignments	• Yes • No
GamingAndFocus	to determine if the games really affect student ability to focus during lectures	• Yes • No
GamingAndPriortization	to determine if the student has missed any meals or any important activities just to play games	• Yes • No
GamingAndPhysicalHealth	to determine if the student experienced any physical pain while playing video games	• Yes • No
GPA_Impact	To determine if Student GPA affected negatively because of video games	• Yes • No

Table 1 Attributes description and values

OBJECTIVE

After collecting and analyzing data using survey and creating dataset. We need to use data mining techniques to identify whether video games have an impact on university students' performance. In this report we'll use two techniques: Decision Tree which is a supervised learning technique used to predict data and apriori which is an unsupervised technique used to find patterns or similarities between different attributes.

First Method: Decision Tree

The first approach we're going to conduct this study is decision tree. Decision tree is a supervised learning algorithm for classification which contains nodes, branches, and leaves. First, we'll load our dataset into Weka; "GPA_Impact" is the class attribute for this study since we want to know if playing video games has any real impact on university student performance. J48 classifier must be chosen to conduct a decision tree and the full training set will be used.

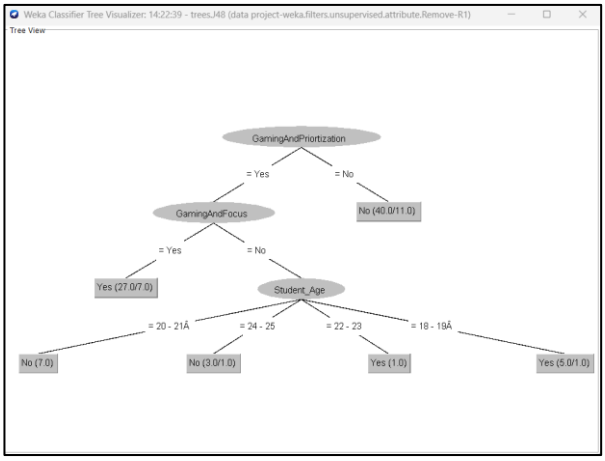

Figure 2 Decision Tree

```
J48 pruned tree
------------------

GamingAndPriortization = Yes
|   GamingAndFocus = Yes: Yes (27.0/7.0)
|   GamingAndFocus = No
|   |   Student_Age = 20 - 21Â : No (7.0)
|   |   Student_Age = 24 - 25: No (3.0/1.0)
|   |   Student_Age = 22 - 23: Yes (1.0)
|   |   Student_Age = 18 - 19Â : Yes (5.0/1.0)
GamingAndPriortization = No: No (40.0/11.0)

Number of Leaves  :     6

Size of the tree :      9
```

Figure 3 Decision Tree

As shown in (figure 2) The decision tree contains 3 nodes with Gaming and Prioritization attribute as its root node, 6 leaves, and a size of 9. Based on the decision tree that we construct we can conclude that if the student did not missed any important activity or meal while playing video games it'll not affect their GPA negatively, but if the student have missed any important activity or meal and he cannot focus during the lecture then it'll affect his GPA negatively for students between ages of 18-19 and 22-23 but it'll not affect the GPA negatively for students between ages of 20-21 and 24-25.

```
=== Summary ===

Correctly Classified Instances          63               75.9036 %
Incorrectly Classified Instances        20               24.0964 %
Kappa statistic                          0.5071
Mean absolute error                      0.3525
Root mean squared error                  0.4198
Relative absolute error                 71.3096 %
Root relative squared error             84.4565 %
Total Number of Instances               83
```

Figure 3 Decision Tree Summary

Figure 3 shows the accuracy of the decision tree, 63 out of 83 are correctly classified with 75.90% and 20 out of 83 are incorrectly classified with 24.09%.

```
=== Confusion Matrix ===

  a  b   <-- classified as
 38  8 |  a = No
 12 25 |  b = Yes
```

Figure 4 Confusion Matrix

Based on figure 4 there are 38 instances of class attribute (GPA_Impact = No) are correctly predicted as No because their GPA will not be negatively affected while 8 instances that are from class attribute (GPA_Impact = No) are incorrectly predicted as Yes.

On the other hand, 25 instances of class attribute (GPA_Impact = Yes) are correctly predicted as Yes because their GPA will be negatively affected while 12 instances of class attribute (GPA_Impact = Yes) are incorrectly predicted as No.

The decision tree was able to predict if the video games really affect university student performance or not but we are still not sure if the result are accurate or not. So, we'll use another data mining technique called Association rule-Apriori to make sure that the results are accurate.

Second Method: Association Rule - Apriori

The Apriori algorithm is a well-known data mining technique that uses data to mine frequent sets of appropriate association rules.

Based on a 90% confidence level, this function will display the relationships between the characteristics and the class as well as the optimum rules for the given data.
From this data collection, we determined 10 rules, as shown below:

```
Minimum support: 0.25 (21 instances)
Minimum metric <confidence>: 0.9
Number of cycles performed: 15

Generated sets of large itemsets:

Size of set of large itemsets L(1): 16

Size of set of large itemsets L(2): 57

Size of set of large itemsets L(3): 41

Size of set of large itemsets L(4): 3
```

Figure5 Apriori

```
Best rules found:

1. GamingVSStudying=No GPA_Impact=No 24 ==> TimeManagementSkills=Yes 24    <conf:(1)> lift:(1.26) lev:(0.06) [4] conv:(4.92)
2. GamingVSStudying=No 33 ==> TimeManagementSkills=Yes 32    <conf:(0.97)> lift:(1.22) lev:(0.07) [5] conv:(3.38)
3. DailyGamingHours=1-2 GamingAndPriortization=No 29 ==> TimeManagementSkills=Yes 28    <conf:(0.97)> lift:(1.21) lev:(0.06) [4] conv:(2.97)
4. GamingAndPriortization=No GPA_Impact=No 29 ==> TimeManagementSkills=Yes 28    <conf:(0.97)> lift:(1.21) lev:(0.06) [4] conv:(2.97)
5. Student_Age=20 - 21Â  GamingAndPriortization=No 23 ==> TimeManagementSkills=Yes 22    <conf:(0.96)> lift:(1.2) lev:(0.04) [3] conv:(2.36)
6. Student_Gender=Female GamingVSStudying=No 23 ==> TimeManagementSkills=Yes 22    <conf:(0.96)> lift:(1.2) lev:(0.04) [3] conv:(2.36)
7. DailyGamingHours=1-2 GamingVSStudying=No 23 ==> TimeManagementSkills=Yes 22    <conf:(0.96)> lift:(1.2) lev:(0.04) [3] conv:(2.36)
8. GamingAndFocus=Yes GamingAndPriortization=No 22 ==> TimeManagementSkills=Yes 21    <conf:(0.95)> lift:(1.2) lev:(0.04) [3] conv:(2.25)
9. DailyGamingHours=1-2 GamingAndPriortization=No GPA_Impact=No 22 ==> TimeManagementSkills=Yes 21    <conf:(0.95)> lift:(1.2) lev:(0.04) [3] conv:(2.25)
10. DailyGamingHours=1-2 GamingAndFocus=Yes 30 ==> TimeManagementSkills=Yes 28    <conf:(0.93)> lift:(1.17) lev:(0.05) [4] conv:(2.05)
```

Figure6 Best rules found.

The confidence level for all the rules is 100%. The rules show the cases where Students demonstrate good time management skills:

1. Students who prioritize studying over gaming and have a good GPA.
2. Students who don't prioritize gaming over studying.
3. Students who spend 1-2 hours on gaming and don't prioritize it.
4. Students who don't prioritize gaming and don't have a negative impact on their GPA.
5. Female students aged 20-21 who don't prioritize gaming
6. Female students who don't prioritize gaming over studying.
7. Students who spend 1-2 hours on gaming and don't prioritize it over studying.
8. Students who prioritize gaming focus and don't prioritize gaming over other tasks.
9. Students who spend 1-2 hours on gaming, don't prioritize it over other tasks, and don't have a negative impact on their GPA.
10. Students who spend 1-2 hours on gaming and prioritize gaming focus.

CONCLUSION

According to the results and outputs obtained from the application of data mining methodologies, it can be concluded: Using Weka as a data mining tool to apply various data mining techniques such as classification, clustering, and correlation helped the study to discover and analyze results from the survey easily and efficiently. From our point of view, playing video games has some detrimental effects on student performance in universities, although the results were very close, and for the majority of students, video games did not affect their performance. But those affected by video games negatively should organize their time between games and study. Also use the time to develop hobbies instead of playing.

ACKNOWLEDGMENT

We would like to express our gratitude to Dr. Ahmed for the amazing opportunity he gave us to expand on the knowledge he had taught us. We also like to thank Miss Hajer for all of her assistance in the lab and the hard work she put in to help us understand the topic. We enjoy working on this project, and it has taught us a lot.

REFERENCES

[1] Twin, A. (2023) What is data mining? how it works, benefits, techniques, and examples, Investopedia. Available at: https://www.investopedia.com/terms/d/datamining.asp (Accessed: 20 May 2023).

[2] Wright.J (2011). "The effects of video game play on academic performance". Modern Psychological Studies [Online]. vol. 17, issue 1. PP 2-6. Available: https://scholar.utc.edu/cgi/viewcontent.cgi?article=1197&context=mps.

[3] Babu.P, Ravindra.P.N, Mudabasappagol.C (2019, December). "Impact of video-games on academic performance and sleep duration in medical students ". Indian Journal of Clinical Anatomy and Physiology [Online]. vol. 6, issue 4. PP 1-4. Available: https://www.ijcap.org/journal-article-file/10406

TEAM DISTRIBUTION

NAME	ID	TASKS
Manar Eyad Ahmed	202006306	1| ATTRIBUTES 2| OBJECTIVES 3| DECISION TREE 4| CONFUSION MATRIX
Walaa Hussain Ali Ahmed	202104424	1| APRIORI
Fatema Hussain Habib	20192408	1| INTRODUCTION 2| SIMILAR WORK 3| ABSTRACT 4| DATA PREPARATION
Zahra Abdulghani Mansoor	20164923	1| DATA COLLECTION 2| CONCLUSION 3|ACKNOWLEDGEMENT