Global Business and Management Research: An International Journal
Vol. 14, No. 1 (2022)

Feature Selection of Digital Addiction Prediction Among Students Using WEKA

Lailatul Faizah Abu Hassan
Faculty of Business and Management, Universiti Teknologi MARA, 42300 Puncak Alam, Malaysia

Nur Atiqah Rochin Demong*
Faculty of Business and Management, Universiti Teknologi MARA, 42300 Puncak Alam, Malaysia
Email: rochin@uitm.edu.my

Mohammad Zaim Mohd Salleh
Faculty of Business and Management, Universiti Teknologi MARA, 42300 Puncak Alam, Malaysia

Emi Normalina Omar
Faculty of Business and Management, Universiti Teknologi MARA, 42300 Puncak Alam, Malaysia

Anisah Alwi
Faculty of Business and Management, Universiti Teknologi MARA, 42300 Puncak Alam, Malaysia

** Corresponding Author*

Abstract

Purpose: This study is aimed to predict the level of digital addiction among students located at selected public university based on selected attributes namely gender, age, education level, marital status as demographic profiles and elements of digital addiction scale inclusive of overuse, non-restraint, inhibiting the flow of life and emotional state have been adopted for the analysis.

Design/methodology/approach: In this study, students from Malaysia's public universities were chosen at random. The study's research design is correlational research, and a set of questionnaires was created specifically for the purpose of the study. The poll was conducted solely at the designated public university, with 128 participants. The relevant data acquired considerably helped to the knowledge of which service elements appear to have significant positive explanatory power on digital addiction among students at selected public colleges. This study investigated and quantified the degrees of digital addiction in 128 respondents based on demographic profiles with four variables as well as digital addiction aspects such as overuse, non-restraint, obstructing the flow of life, and emotional state. This study predicts the digital addiction level based on an established theoretical framework for lists of items to measure the four dimensions listed.

Findings: The feature selection analysis was analyzed using the WEKA tool on a dataset of 128 instances and 8 determinants. On the WEKA interface, based on feature selection using InfoGainAttributeEval, it is predicted that emotional state ranked first followed by inhibiting the flow of life, overuse and non-restraint. The attribute evaluator for CorrelationAttributeEval generate similar results as InfoGainAttributeEval for four elements of digital addiction scales.

Research limitations/implications: The authors used public universities' students as sample by employing a quantitative research design. For future research, the author suggested using qualitative approaches such as interviews and observation. The researchers may be able to gather more exact and reliable data that reflects personal ideas and impressions by having direct conversations with the respondents.
Practical implications: Internet addiction is a serious problem that should be diagnosed (and, if required, treated) in maturity, dispelling the myth that this form of addiction only occurs during periods of life when people lack free will, such as adolescence or youth.
Originality/value: A timely review of the internet abuse/addiction phenomenon is presented with the objective of increasing awareness, debate, and additional empirical research which could contribute to the early prevention of internet addiction problems.

Keywords: Data Mining, Digital Addiction, Feature Selection, Prediction

Introduction

The Internet has become an essential component of practically every college student's daily routine, and most of them appear to benefit from using the Internet to conduct research and communicate with teachers and peers (Jones, Johnson-Yale, Perez, & Schuler, 2007 as described in O'Brien, 2011). According to Kesici and Tunc (2018), addiction is described as an impulse depending on a habit of a certain activity or substance use, despite the fact that it has long-term negative repercussions on the individual's physical, social, emotional, and mental health, as well as his/her financial circumstances (Young, Yue & Ying, 2011). Students' internet addiction should be treated seriously before it becomes an issue. The responsibilities of parents and educators in assessing the possibility of addiction in their children and students should be highlighted. For treating addicts, psychological tactics such as motivational interviewing must be designed specifically for Internet addiction. Researchers have yet to investigate how the use of digital technologies may affect academic achievement. This study investigated how digital technologies and other Internet use could affect student performance in college courses by questioning participants about characteristics such as attendance, procrastination, distraction, time spent studying, access to technology, and use of the Internet for research purposes. More research into students' attitudes about digital technologies and their impact on social life is needed to preserve societal peace and happiness. This study is specifically designed to learn more about students' perceptions of the academic impact of their digital technologies and other Internet use.

Literature Review

Internet Use

Internet use is an important aspect of everyone's everyday lives, especially for college students who utilise it for social interaction, entertainment, and education. Internet use has become the most popular recreational and academic activity for university students as a result of rapid technological advancement. Internet addiction causes social and psychological problems such as depression, loneliness, low self-esteem, and life satisfaction, as well as other mental health issues, and has a bigger influence on their academic performance (Subhashini & Praveen, 2018). The purpose of this study is to determine the relationship between the internet and depression, as well as the role of gender in online addiction.

Technology and Internet use, particularly among college students, have been linked to increased communication with friends and family members (Clark & Frith, 2005). Numerous studies have found that the primary usage of the Internet for college students is interpersonal contact via email, instant messaging, and chat services (Hampton & Wellman, 2001; Howard

et al., 2001; Kraut et al., 1998; McKenna and Bargh, 2000; Jones, 2002). Despite the fact that most people agreed on the benefits of the Internet as a healthy productive activity, such as students catching up on the opportunity by visiting websites, engaging in chat-rooms, wonderful tool for research, and other (Baturay & Toker, 2019).

Digital Addiction

Many students, on the other hand, fall behind in their studies as a result of excessive investments in online interactions, a condition known as internet addiction (Hansen, 2002). Many studies have suggested that people can become addicted to the Internet, resulting in bad social behaviour, habits, and talents (Young, 1996; Scherer, 1997; Kraut et al., 2002; Kubey et al., 2001; Nalwa & , 2003; Baturay & Toker, 2019; Shehata & Abdeldaim, 2021).

Goldberg (1996) coined the term "internet addiction" to describe pathological obsessive internet use (Ranganatha & Usha, 2017). In the literature, it is referred to as pathological internet usage, problematic internet use, compulsive online use, and internet overuse. Those suffering from internet addiction, like those suffering from other addictions, use the delusory world to communicate with actual people over the Internet as a substitute for real-life human connection, which they are unable to obtain properly (Suresh, 2016; Shehata & Abdeldaim, 2021).

Digital Addiction and Performance

The internet has become ingrained in the lives of students. Many students used the internet primarily for educational purposes; however, some students squandered their time by visiting improper, non-educational websites. According to a study, students' academic progress is impacted by how they use the internet, whether for educational or non-educational purposes (Li et al., 2014). Adolescents who use the internet extensively join chat rooms, which causes them to stay up late and affects their focus and attention in class, resulting in a decline in their academic performance (Leung & Lee, 2012; Usman, Alavi, & Shafex, 2013; Najmi et al, 2014; Kir & Sulak, 2014, Rashid et al, 2020).

Other indicators of emotional instability include stress, pressure, or tension. According to one study, stress has a negative relationship with academic performance, meaning that the higher the stress level, the worse the academic performance (Samaha & Hawi, 2016). Students' stress levels increased when they spent too much time on the internet and had little time to study. This finding is backed up by Khan, Altaf, and Kausar (2013), as well as previous research by Safree, Yasin, and Dzulkifli (2009), who found that depression, anxiety, and stress are negatively associated to academic performance.

Digital Addiction Scale

As shown in Figure 1, the criteria for process and output based on the selected theoretical model by Kesici and Tunç (2018) to assess the degree of digital addiction among students in selected public university are based on overuse, non-restraint, inhibiting the flow of life and emotional state components.

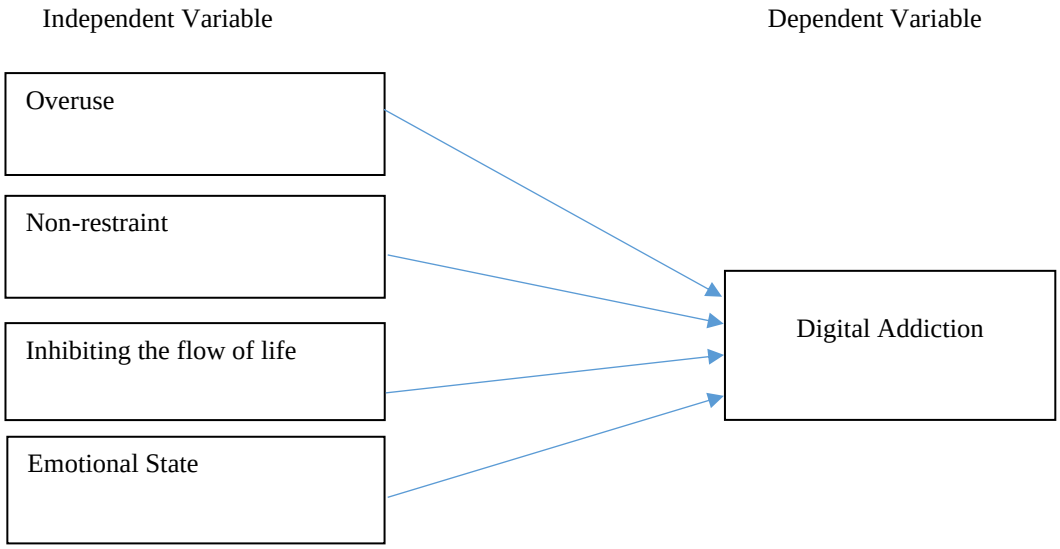

Figure 1: Digital Addiction Scale Framework adopted from Kesici and Tunç (2018).

Design

The purpose of this research was to predict the level of digital addiction among university students. The authors use a data mining approach in their research. The field of information technology is always evolving and increasing. The technique of collecting meaningful knowledge from huge amounts of data is known as data mining.

As a result, correlational research is chosen as the research design. The relationship between two variables is described using this method. Correlational research, according to Sekaran and Bougie (2013), is conducted when the researcher is interested in depicting the factors that are related to the issue.

According to Crossman (2011), the unit of analysis refers to the major entity that is analysed in the study. Based on another assertion by Sekaran & Bougie (2013), the unit of analysis refers to the level of aggregation of the data collection during the subsequent data analysis stage. Hence, the unit of analysis that was used in this study is individual that refers to universities' students addicted to digital environment. Their individual responses on the studied variables were used to generate the results of the feature selection and Apriori algoritm.

This study examined a total of 128 usable questionnaires. This study's main instrument was a questionnaire that had been adapted and modified from previous studies. The questionnaire was specifically designed to match the nature of the research and to answer the study's research questions. All respondents received the questionnaire in person.

Figure 2 depicts the research performed on the collected dataset, indicating that pre-processing procedures such as attribute discretization to nominal data type and attribute selection strategy were used prior to employing the algorithm. Attribute selection is done to prioritize the attributes in order to find the most important components.

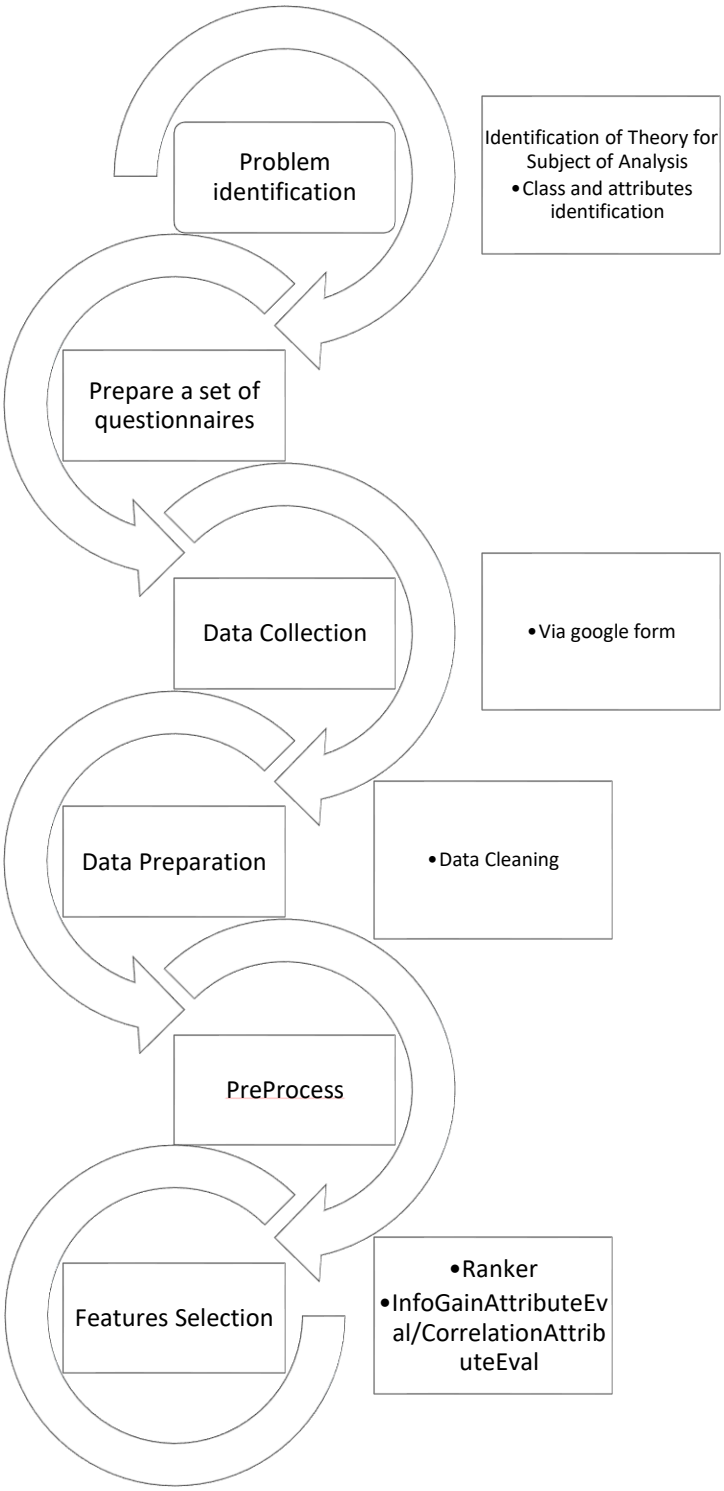

Figure 2: Research Design

Analysis and Results

A total of 200 questionnaires were delivered to Malaysian public university students who are actively and addicted to the usage of the internet. As stated in Table 1, a total of 128 suitable were evaluated.

Table 1 Attributes and data description

Demographic Variable	Data Type	Frequency	Percentage (%)
Gender	Nominal		
Male		32	25
Female		96	75
Education Level	Nominal		
Diploma		5	3.9
Bachelor Degree		110	85.9
Master Degree		11	8.6
PhD		2	1.6
Marital Status			
Single		120	93.75
Married		8	6.25
Age	Nominal		
Below 30		118	92.19
31 to 41		7	5.47
More than 41		3	2.34
Overuse	Nominal		
Strongly Disagree		0	0
Disagree		24	18.75
Neither Agree nor Disagree		57	44.53
Agree		40	31.25
Strongly Agree		7	5.47
Non-restraint	Nominal		
Strongly Disagree		11	8.59
Disagree		35	27.34
Neither Agree nor Disagree		51	39.84
Agree		25	19.53
Strongly Agree		6	4.70
Inhibiting the flow of life	Nominal		
Strongly Disagree		18	14.06
Disagree		33	25.78
Neither Agree nor Disagree		53	41.41
Agree		19	14.84
Strongly Agree		5	3.91
Emotional State	Nominal		
Strongly Disagree		2	1.56
Disagree		25	19.53
Neither Agree nor Disagree		66	51.56
Agree		31	24.22
Strongly Agree		4	3.13
Class (Digital Addiction Level)	Nominal		
Very Low		2	1.56
Low		4	3.13
Moderate		40	31.25
High		**57**	44.53
Very High		25	19.53

Pre-process Results

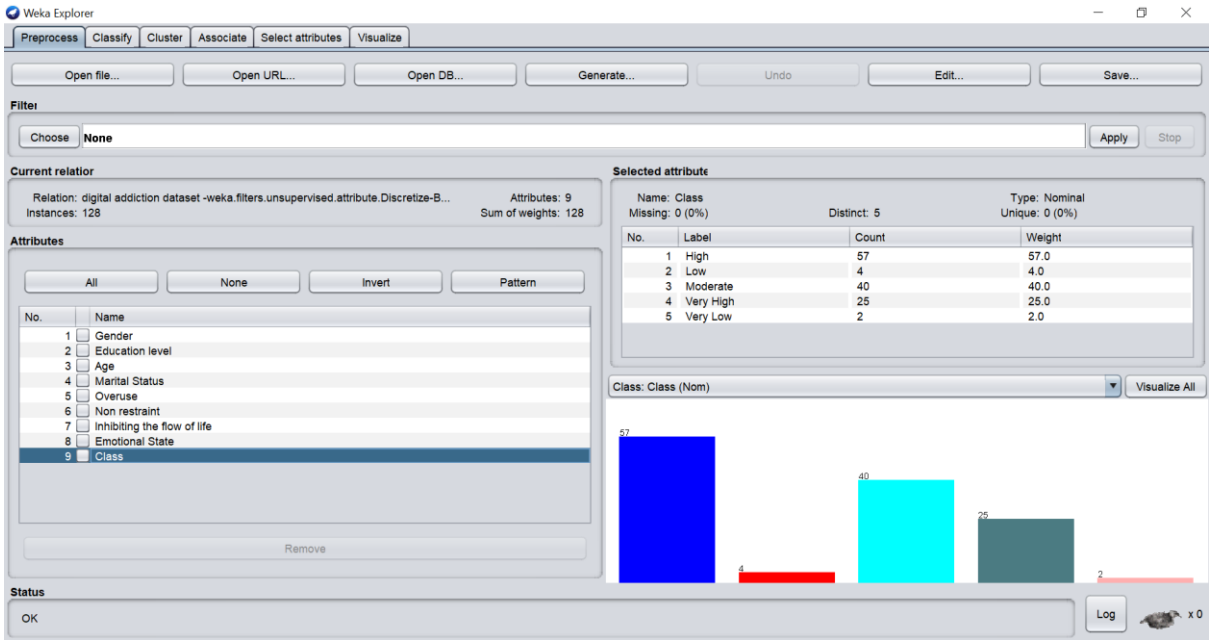

Figure 3: List of attributes and class for the digital addiction analysis.

Figure 3 depicted the class of the analysis which refer to the level of digital addiction among student in selected public universities. Based on Table 1 and Figure 3, there are five levels of digital addiction namely very low, low, moderate, high and very high. The result shows that it is dominated by blue color which represent high level of digital addiction with the value of 57 followed by moderate (40), very high 925, low (4) and very low (2). This study is confirmed by Rashid, Aziz, Rahman, Saad, and Ahmad (2020), who conducted a study with secondary school students and discovered that the students have their own reasons for their dependence on mobile phones, which leads to addiction and the inability to live without them.

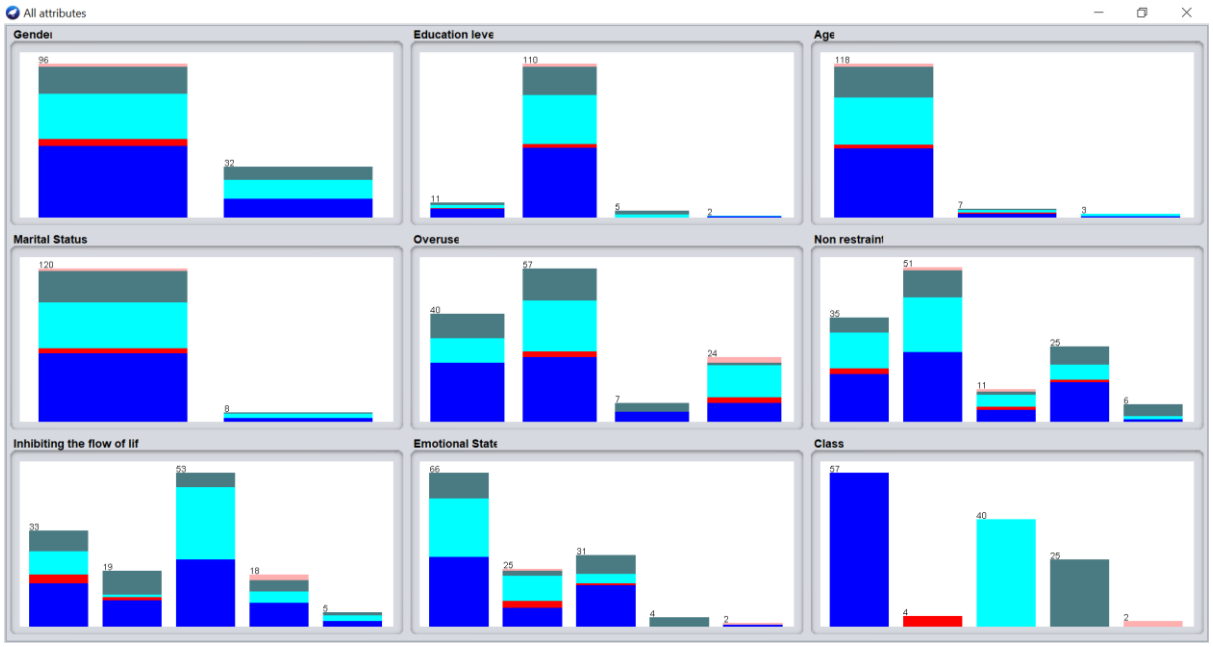

Figure 4: Visualize All Results

Figure 4 depicts the eight qualities or drivers of digital addiction among students, with blue representing the agree degree of satisfaction and cyan representing the highly agree level. The colour red signifies a high level of digital addiction. It is possible that the pupils became glued to digital worlds. Figure 4 contains the bar labelling and data seen in Table 1. Female students, for example, have a higher level than male students since a larger proportion of them and are under 30 years old and typically hold a bachelor's degree. Furthermore, the majority of responders are still single.

Feature Selection

The feature selection algorithm in WEKA determines the most significant qualities by employing correlation-based attribute assessment, as illustrated in Figure 5 and Figure 6.

InfoGainAttributeEval

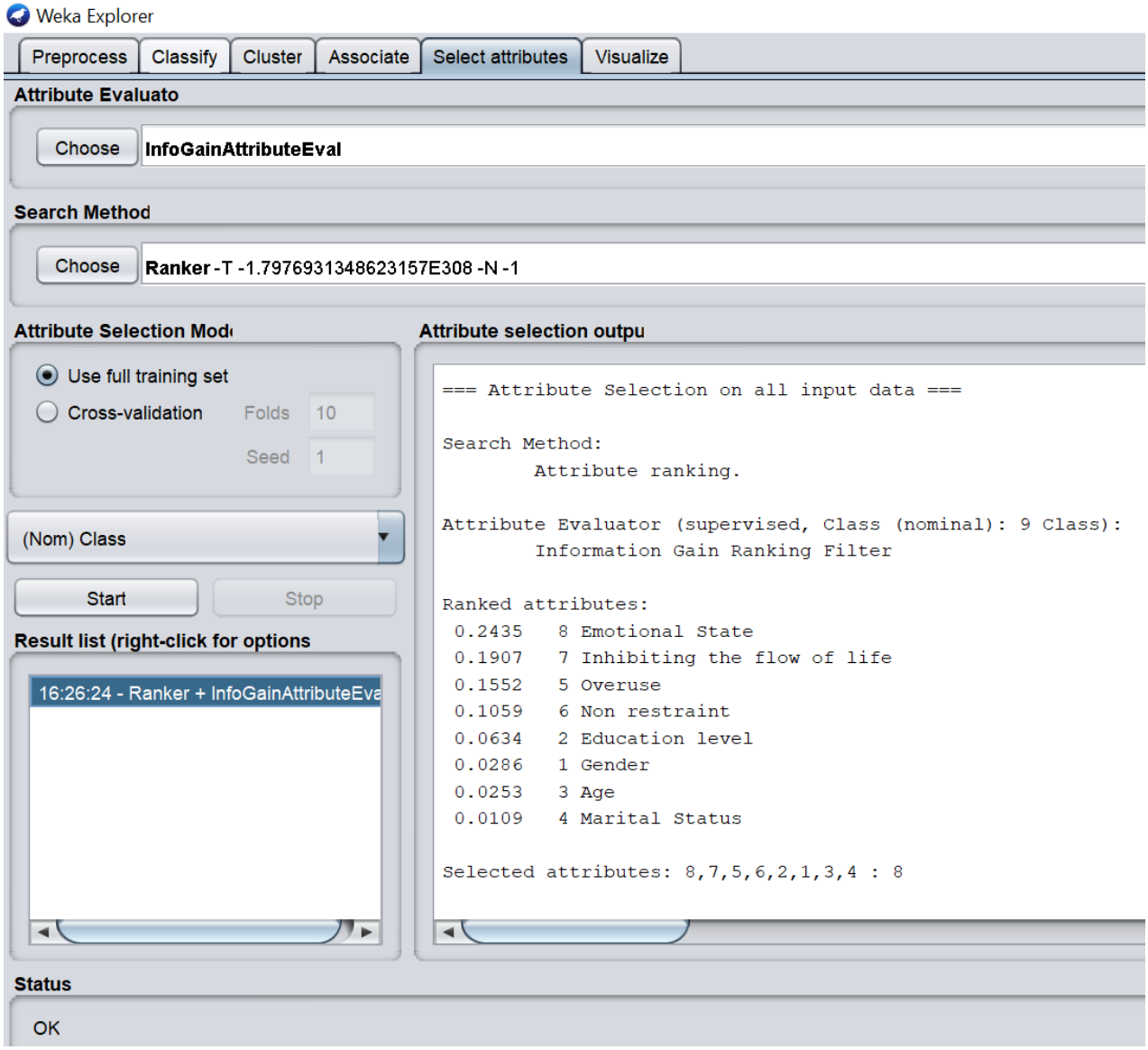

Figure 5: InfoGainAttributeEval Result

InfoGainAttributeEvaluation evaluates the worth of an attribute by measuring the information gain with respect to the class. From the result as shown in Figure 5, it can be summarized that the most influential attribute contributing to the level of digital addiction among universities

students is emotional state (Attribute No. 8) followed by inhibiting the flow of life, overuse and non-restraint element which rank top four of the results showing these four elements does have significant affect toward digital addiction level among students. The findings are confirmed by Rathi and Budhiraja's (2019) study of undergraduate students and the effects of social networking sites on academic performance, which indicated a favourable association between the use of social media networking and the academic performance of undergraduate students. For demographic profiles, it is leaded by educational level, followed by gender, age and marital status. The degree of digital addiction among public university students can be measured by their reliance on digital devices for them to express their emotion when dealing with daily activities neither formal or informal events. Thus, this led to the higher percentage and ranked as main determinant towards digital addiction level.

CorrelationAttributeEval

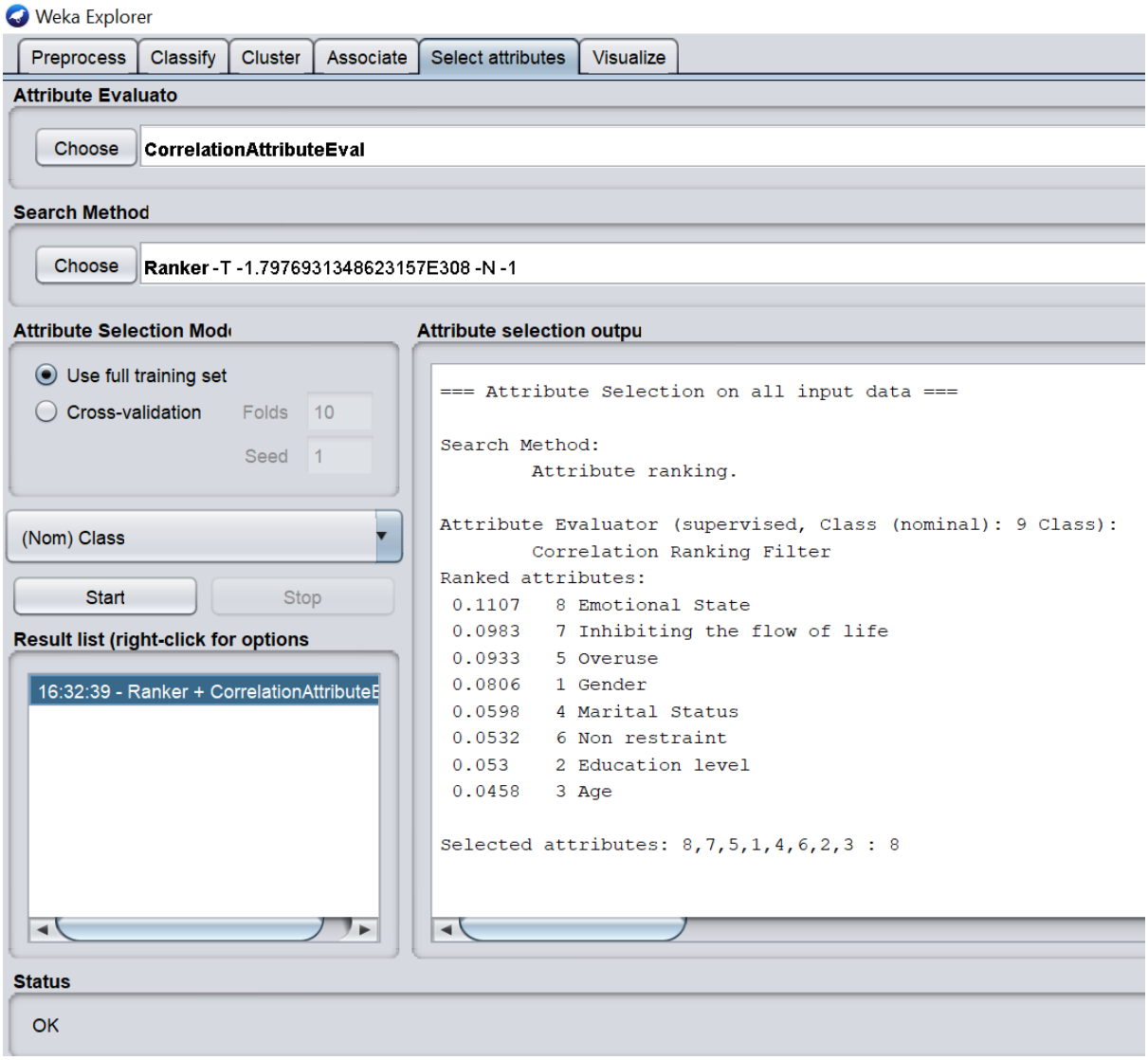

Figure 6: CorrelationAttributeEval Result

CorrelationAttributeEvaluation ealuates the worth of an attribute by measuring the correlation (Pearson's) between it and the class. As shown in Figure 6, the results is quite similar with InfoGainAttributeEval for the first three attributes that affect the digital addiction among

students which dominated by emotional state followed by inhibiting the flow of life and overuse as compared to demographic profiles. Thus, it can be concluded that the four elements adopted for this study does affect the level of digital addiction level among students.

Association Rules Results (Apriori Algorithm)

Data is analysed using association rules to find common if/then hidden knowledge. The key relationships are discovered through the use of support and trust criteria. The Association regulation is made up of two parts. The Apriori algorithm is the most often used method for gathering correlation-based data (Nafie & Hamed, 2018). According to the results, the lowest level of support was 0.8 (80%) (102 occurrences), the metric (confidence) was 0.9 (90%), and there were 4 cycles. Figure 7 depicts the author's selection of the finest guidelines. The Apriori algorithm's default setting for the number of best rules generated is ten, and the results show that education level and age dominated with the most best rules generated, from the ten rules produced.

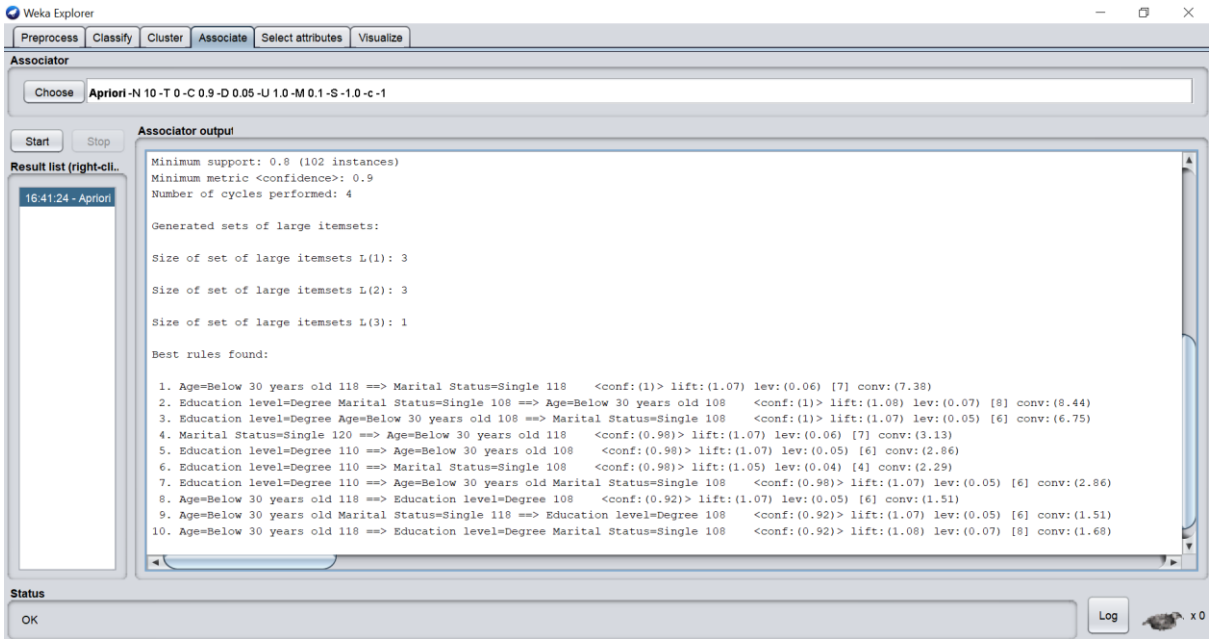

Figure 7: Apriori Result

From the result shown in Figure 7, education level and age have high impacts towards digital addiction among students with the valued of age is below 30 years old and level of education is degree level with the marital status is single.

Discussion and Conclusion

Excessive Internet use may have a negative impact on students' attendance, academic performance, and social interactions in the long run. Previous research has demonstrated that the Internet can divert students' attention away from their studies (Chang & Law, 2008; Tsai et al., 2009; Nalwa & Anand, 2003). Furthermore, if students continue to participate in excessive Internet use, especially those interested in cyber-relationships, a lack of social skills with real human beings may be perceived as an issue. Social networking's social support and anonymity can lead to more severe addicted behaviour (Suresh, 2016). The institution can also raise student knowledge about the detrimental impacts of internet addiction by holding campaigns, flyers, and seminars, which could contribute to the early prevention of internet addiction problems.

An in-depth study that examined the impact of Internet use on young people's academic performance and social lives is also seen to be useful in delivering conclusive findings and recommendations.

Furthermore, rather than employing a quantitative research design, future researchers are advised to use a qualitative design approach, which was used in this study. Future researchers will also be required to perform their studies using qualitative data collection approaches such as interviews and observation. The researchers may be able to gather more exact and reliable data that actually reflects personal ideas and impressions by having direct conversations with the respondents.

Acknowledgement

We would like to thank Faculty of Business and Management, Universiti Teknologi MARA for awarded us the research grant [600-IRMI 5/3/DDF (FPP) (009/2019)].

References

Baturay, M.H. & Toker, S. (2019). Internet addiction among college students: Some causes and effects. Educ Inf Technol 24, 2863–2885. https://doi.org/10.1007/s10639-019-09894-3

Chang, M.K., and Law, S.P.M. 2008, Factor structure for Young's Internet addiction test: A confirmatory study. Computers in Human Behavior, 24, 2597-2619.

Chen L, Wang L, Qiu XH, Yang XX, Qiao ZX, Yang YJ, et al. Depression among Chinese university students: prevalence and sociodemographic correlates. PloS one. 2013; 8(3): e58379

Clark, D. J., & Frith, K. H. (2005). The Development and Initial Testing of the Internet Consequences Scales (ICONS). CIN: Computers, Informatics, Nursing, 23(5), 285–291. doi:10.1097/00024665-200509000-00013

Goldberg, I. (1996) Internet Addiction. Accessed 15 November 2018. Available from: http://web.urz.uniheidelberg.de/Netzdienste/anleitung/wwwtips/8/addict.html

Gününç, S., & Kayri, M. (2010). The profile of internet dependency in Turkey and development of internet addiction scale: Study of validity & reliability. Hacettepe University Journal of Education, 39(39): 220-232.

Hampton K, Wellman B. (2001). Long distance community in the network society. Am. Behav. Sci. 45:476–95

Howard PEN, Rainie L, Jones S. (2001). Days andnightsontheInternet.Am.Behav.Sci.45: 383–404

Jones, S., Johnson-yale, C., Pérez, F. S., & Schuler, J. (2007). The Internet Landscape in College. Yearbook of the National Society for the Study of Education, 106(2), 39–51. doi:10.1111/j.1744-7984.2007.00114.x

Kakaraki, S., Tselios, N. & Katsanos, C., (2017). Internet addiction, academic performance and personality traits: A correlational study among female university students, International Journal of Learning Technology, 12(2).

Kesici, A., & Tunç, N. (2018). The Development of the Digital Addiction Scale for the University Students: Reliability and Validity Study. *Universal Journal of Educational Research, 6*, 91-98. https://doi.org/10.13189/ujer.2018.060108

Kır, İ., & Sulak, Ş. (2014). Determining internet addiction level of faculty of education students. Electronic Journal of Social Sciences, 13(51), 150-167.

Leung, L., & Lee, P. S. N. (2012). Internet Addiction Symptoms, and Internet Activities on Academic Performance. Social Science Computer Review, 30, 403-418. http://dx.doi.org/10.1177/0894439311435217

McKenna, KYA, & BarghJA. (2000). Coming out in the age of the Internet: identity 'demarginalization' through virtual group participation. J. Personal. Soc. Psychol. 75:681–94

Najmi Hayati Usman, Masoumeh Alavi, Syed Mohamed Shafeq (2014). Relationship between Internet Addiction and Academic Performance among Foreign Undergraduate Students,Procedia - Social and Behavioral Sciences,114, 845-851, https://doi.org/10.1016/j.sbspro.2013.12.795.

Nalwa, K., and Anand, A.P. (2003). Internet addiction in students: a cause of concern. Cyberpsychology & Behavior,6, 653-656

Nafie, F., & Hamed, A. (2018). Usage Apriori and clustering algorithms in WEKA tools to mining dataset of traffic accidents. *Journal of Information and Telecommunication, 2*, 1-15. https://doi.org/10.1080/24751839.2018.1448205

Ranganatha SC & Usha S. (2017). Prevalence and pattern of internet addiction among medical students, Bengaluru. International J Community Med Public Health. 4(12):4680-4.

Rashid, J. A., Aziz, A. A., Rahman, A. A., Saaid, S. A., & Ahmad, Z. (2020). The Influence of Mobile Phone Addiction on Academic Performance Among Teenagers. Jurnal Komunikasi: Malaysian Journal of Communication, 36(3).

Rathi, A., & Budhiraja, A. (2019). Effects of social networking sites on the academic performance of undergraduate students. ABS International Journal of Management, 7(1), 69-74.

Sekaran, U. and Bougie, R. (2013) Research Methods for Business: A Skill-Building Approach. 6th Edition, Wiley, New York

Shehata, W.M., Abdeldaim, D.E. (2021). Internet addiction among medical and non-medical students during COVID-19 pandemic, Tanta University, Egypt. Environ Sci Pollut Res. https://doi.org/10.1007/s11356-021-14961-9

Subhashini & Praveen (2018). An era of digital slavery: a study on internet addiction among professional college students of Hassan, Karnataka. International Journal Community Med Public Health, 5(7), 2977-2983

Suresh Prabu P. (2016). Internet Addiction Among Arts and Science College Students. IOSR J Humanities Social Sci., 21(9):76-81.

Tsai, H. F., Cheng, S. H., Yeh, T. L., Shih, C. C., Chen, K. C., Yang, Y. C., & Yang, Y. K. (2009). The risk factors of Internet addiction-A survey of university freshmen. Psychiatric Research, (167), 294-299.

Usman, N. H., Alavi, M., & Shafex, S. M. (2013). Relationship between Internet Addiction and Academic Performance among Foreign Undergraduate Students, Procedia – Social And Behavioral Sciences, 114, 845-851.

You, S., Kim, E., & Lee, D. (2017). Virtually real: exploring avatar identification in game addiction among massively multiplayer online role-playing games (MMORPG) players. Games and Culture, 12(1), 56-71.

Young, K. S., Yue, X. D., & Ying, L. (2011). Prevalence estimates and etiologic models of Internet addiction. Internet addiction: A handbook and guide to evaluation and treatment, 3-17.