Internet Usage: significant factors affecting students

Rana Jyoti Chakma[1*], Sayed Asaduzzaman[1], Tanjim Mahmud[1], and Hasin Rehana[2]

[1]Department of Computer Science & Engineering,
Rangamati Science and Technology University, Rangamati, Bangladesh.

[2]Department of Computer Science & Engineering,
Rajshahi University of and Technology, Rajshahi, Bangladesh.

Abstract

Currently, the Internet has become an integral part of our daily lives and its use in education is required more than ever. In this research, a data mining approach has been used to determine the key factors affecting students due to Internet usage.

In this research, data were collected from the students at all the levels of Rangamati Science and Technology University (RMSTU) through a structured questionnaire survey method which was designed from background studies. PCA, SVM, Ranker with CFP attribute evaluator, Association analysis took place in this research. The findings of this research revealed that factors like social media use, internet problems, addiction to online games, etc are affecting students and their academic performance. The research showed that the impact of Internet usage on university students is high, and it impacts mostly junior male students. More data with extended parameters that can be used with future data mining analysis will express more significant patterns in the future.

This study recommends using a larger dataset in the future to find more accurate analysis and embarking on remedial programs for the students who are already adversely affected due to the addition of Internet usage.

Keywords: Students, Academic, Performance, Internet, Data Mining, Internet usage, Factors

Introduction

The Internet has now become the main tool and is playing a central role in any academic institution. Students are now spending more time on the Internet and many of them are changing their text and reference books with online editions (Senthil, 2018).

It has been reported in various media that students are using social network media, such as Facebook, IMO, Instagram, etc., and other video sharing networks, such as YouTube, Vimeo, Daily Motion, etc., much more rate than ever before. A study by (Al-Jubayer, 2013) indicated that the use of Facebook was very important for 25% of the respondents, and half of the respondents used Facebook on a fairly regular basis. Forty-six percent of the teenagers in the study indicated the time they were most likely online was between 10 p.m. to 12 a.m. While this has benefitted many students in terms of getting access to free online learning resources, research findings have shown that excessive Internet usage adversely affects one's physical health, family life, and academic performance (Akhter, 2013). Academic related problems include a decline in study habits, a drastic drop in grades, missing classes, and poor integration in extracurricular activities.

In light of the above background, thus a study on the Impact of Internet Usage on students' academic performance, and their risky online behaviors became very important in Rangamati and around Bangladesh.

The main objective of this research study is to identify the significant factors that are affecting students during their studies. A survey-based study has been taken place in this research using a data mining approach to find out the patterns and factors affecting students. The (Stanislav & Vassu-Dimov, 2009) research analyzed a dataset to find out the significant patterns and factors that tend to make decision trees. All the analysis and final decision tree in this research study revealed that some factors like social media use, internet problems, addiction to online games, etc are affecting students and their academic performance. It also showed that the junior level university students with the age group of 16-25 have the highest impact basis on the risk factors of ranking and machine learning analysis.

Literature Survey

The study by (Emeka & Nyeche, 2016) has shown that the Internet is one of the beneficial tools in this era of digital technology, not only for business but for academic purposes as well, as it greatly enhances the skill and capability of students. A study which was carried on the medical students has revealed that medical students with high internet usage are associated with higher academic performance (Siraj, et al., 2015). Another study by (Nugent, Shannon, McNamee, & Molyneaux, 2015) has found out that a significant proportion of the students, particularly those who spend no time using a computer/laptop for homework underperformed relative to their peers in their GCSE exams.

The research in Indonesia (Puspita & Rohedi, 2018) showed that the usage of the Internet is more likely to create a negative impact on the academic performance of vocational students because it has made the students to social media, online gaming, watching videos on YouTube. The higher usage of the Internet has also been found to minimize the available time for students' social interaction (Soegoto & Tjokroadiponto, 2018). Internet Usage has also created enormous risks to students for being exposed to various threats. These risks and threats have been established through several research studies in other countries. One of them is the EU Kids Online project (Hasebrink, Livingstone, Haddon, & Ólafsson, 2009) which categorizes the risks associated with Internet usage into four categories: commercial risks, aggressive risks, sexual risks, and values risks. The findings of the researchers in Turkey (Yilmaz, Yilmaz, Ozturk, & Karademir, 2017) indicate that the majority of the 2029 students studying in secondary schools have insufficient knowledge regarding information security and computer usage awareness and they could be at risk in online settings towards the threats. A study by (Guan & Subrahmanyam, 2009) has shown that online risks such as addiction, cyberbullying, and sexual solicitation are associated with negative consequences for youth.

The research study by V. Senthil (Senthil, 2018) has concluded that there was no direct relationship between Internet usage with their CGPA.

One, thus, can easily find from the previous literature that some researchers recommend the use of the Internet in academic studies for educational benefits, on one hand; some researchers caution against it due to some negative impacts, on the other hand.

Methodology

This research was conducted by collecting data from all the levels of students at Rangamati Science and Technology University (RMSTU) which has about 438 students.

Data Collection

The population for this research is the entire undergraduate students of Rangamati Science and Technology University (RMSTU), which is about 438 students. As per Krejcie and Morgan's (Krejcie & Morgan, 1970) table for determining a sample size from a given population, it was decided to have a sample size of 205 from different age groups and different classes. All the target 205 students were

approached for data collection, however, only 142 students responded to the survey. Each question in the survey had a 5-point interval scale ranging from strongly disagree (1) to strongly agree (5), which allowed a respondent to answer a question by choosing a score out of 5. Data collected from all the respondents were analyzed to achieve the objectives of this study.

Data Pre-processing Using WEKA

Data in the real-world is incomplete, inconsistent, and noisy. Thus Data cleaning, Data integration, Data transformation, Data reduction, Data discretization are necessary to fill the missing values, detect the conflict of data, and combine data and many others.

WEKA is used for data preprocessing which is one of the vastly used data mining tools now-a-days. WEKA, a data mining tool, is a collection of machine learning algorithms and it contains tools for data pre-processing, classification, regression, clustering, association rules, and visualization. Data preprocessing in WEKA includes three steps they are-load data, preprocess data, analyze attributes. The workflow of WEKA is the followings:

Data → Pre-processing → Data Mining →Knowledge

There are four applications of WEKA in which data mining can be done. They are Explorer, Experimenter, Knowledge Flow, Simple CLI (Command Line Interface). Here Explorer application is used for data preprocessing. Explorer is an environment that supports data preprocessing, attribute selection, learning, and visualization for exploring data. WEKA supports various file formats like-ARFF, CSV, C4.5, binary. Here for analyzing data CSV file format is used for loading a dataset into the WEKA tool. For constructing the CSV file some data were reduced to avoid a clash between data. When the CSV files are containing real prepared data then WEKA tools are used for classification.

Some filtering operations were done on the data which were discrete for a preprocessing purpose. The approaches are used for selecting attributes (Dimitriu, 2015), which are:-

Significant Factors and Prediction Tools

Finding a significant pattern is a very important task as it classifies and makes a grouping of all the parameters. After data preprocessing significant pattern is calculated by some previous study and the results of the Ranker, Greedy Stepwise, and Best First technique are analyzed in WEKA data mining software.

$$SFP = (\sum Wi \times Si)$$

SFP denotes the significant frequent pattern **W** is the weightage value for each parameter and **S** is the frequency of each of the parameters.

Risk Factor Analysis using Orange

Orange is a data mining tool that analyzes data through backend algorithms. PCA, Decision Tree, SVM analysis can be done using Orange. Orange has been used in researches (Dimitriu, 2015), (Assaduzzaman & Al Masud, 2018) to analyze data in finding the significant factors among a dataset.

Results and Discussion

Education Levels

The respondents were asked about their education levels. Fig.-1 shows the major education levels which are- 1st Year (44.45%), the 2nd Year (14.8%), the 3rd Year (12.7%), and the 4th Year (16.9%) students at RMSTU. Also, though insignificant, the respondents included students studying MBA and M.Sc. Engineering in Computer Science and Engineering (CSE).

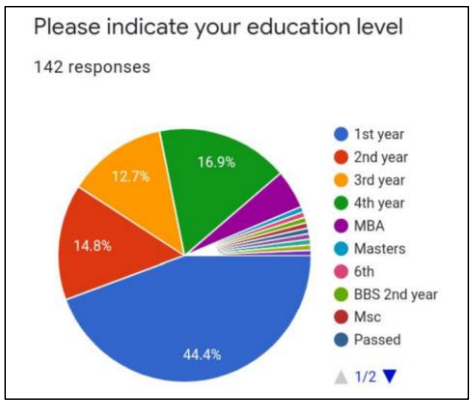

Fig.-1: Education Levels (Years)

Different level students vs Age groups

Box plot analysis of the different year students has been shown in Fig 2 and Fig 3. The box plot shows that for the age group of 21-25 years, the students in 1st, 2nd, 3rd, and 4th years have almost all the same distribution, however, for the age group 16-20 years, the 1st year students have the highest distribution, while the 2nd year students have the second-highest field distribution. Fig.-3 shows that the impact of Internet Usage is higher among male students in all the different education levels.

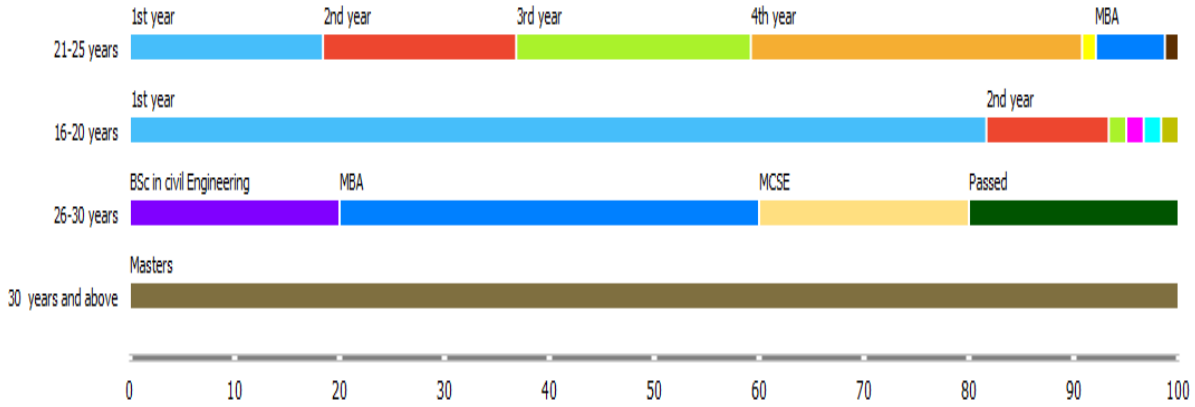

Fig.-2: Box Plot Analysis of the different year students vs age groups

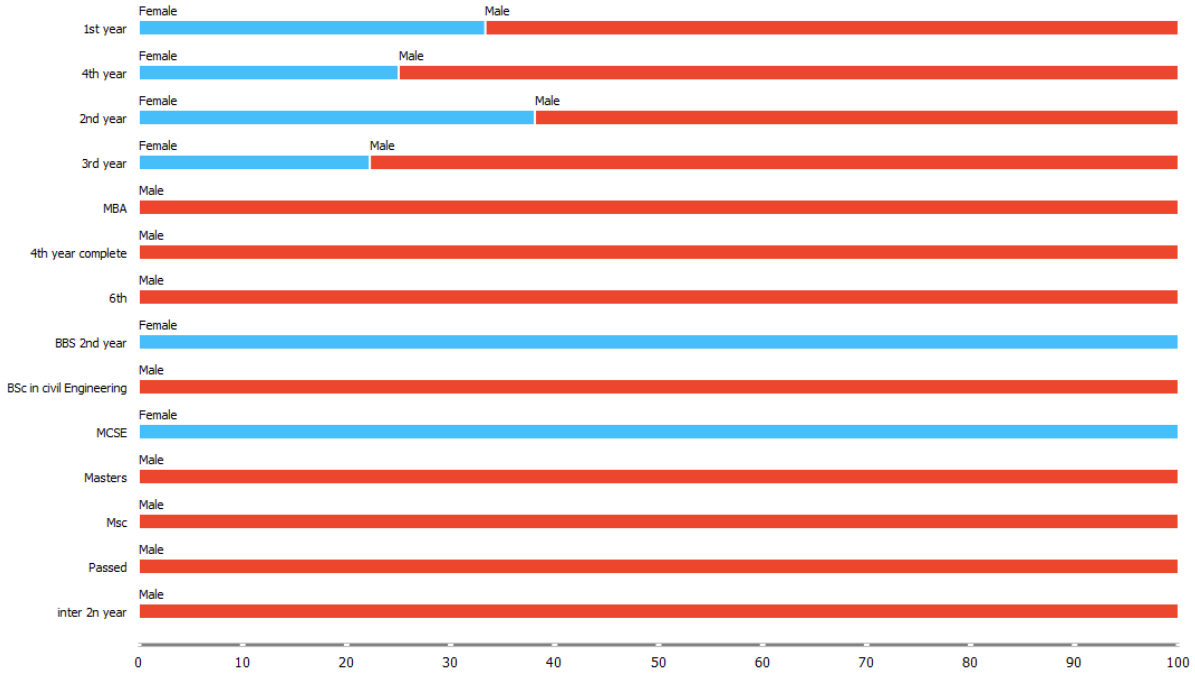

Fig.-3: Box plot analysis with different education level students according to Gender

The probability distribution of students according to Euclidean distances has been illustrated in Fig. 4. PCA curve of the dataset has been depicted in Fig 5. All the results of Fig.-4 and Fig.-5 illustrate the probability of the students in 1^{st}. 2^{nd}, $3^{rd,}$ and 4^{th} years have a proper distribution from 0.1 to 0.

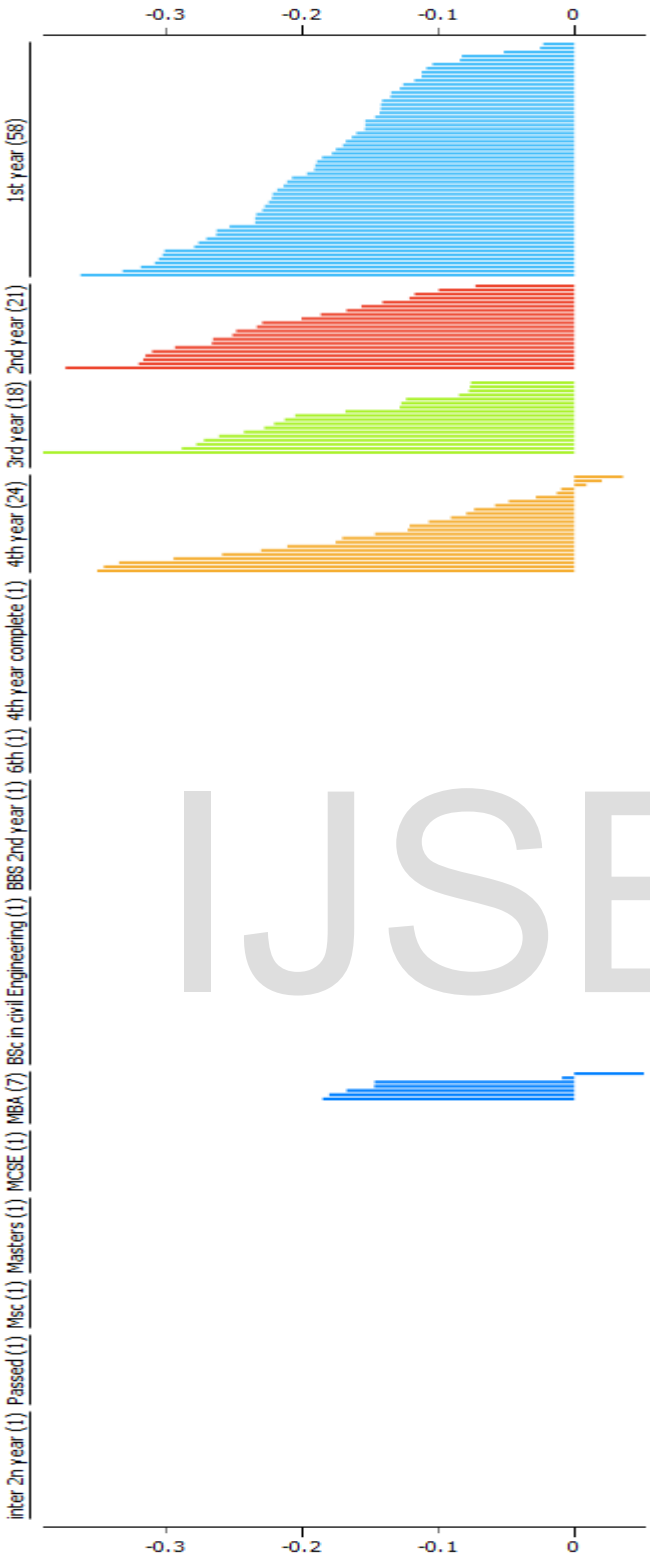

Fig.-4: Probability Distribution according to Eucleadian Distance among the Students

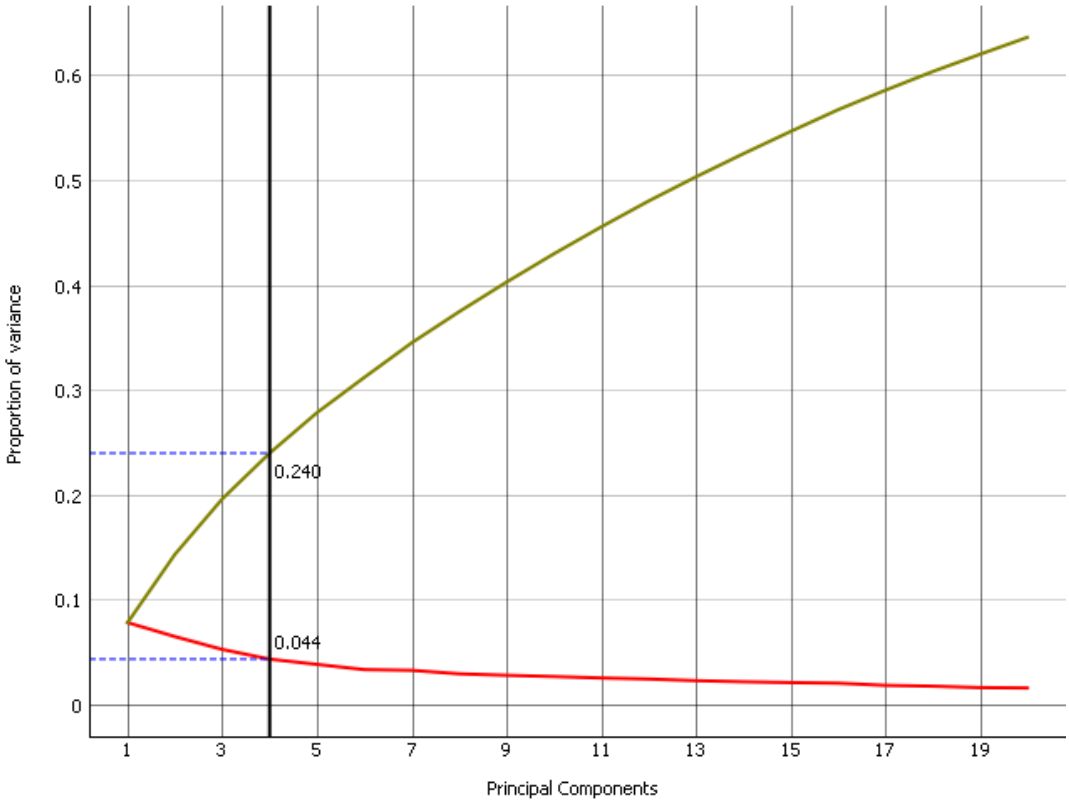

Fig.-5: Principle Component Analysis (PCA) among the dataset

Association and Ranking among the significant factors

Table 1 shows the association among the factors: the use of social networking media and academic performance.

Table 1: Association among the significant factors

1. Do you use the internet for your Academic Studies? =Yes 130 ==> Do you use the internet for Social Networking (Facebook, Messenger etc.)? =Yes 130 <conf:(1)> lift:(1) lev:(0) [0] conv:(0)
2. Do you use the internet for other things? =Yes 130 ==> Do you use the internet for Social Networking (Facebook, Messenger etc.)? =Yes 130 <conf:(1)> lift:(1) lev:(0) [0] conv:(0)
3. Do you use the internet for your Academic Studies? =Yes Do you use the internet for other things? =Yes 120 ==> Do you use the internet for Social Networking (Facebook, Messenger, etc.)? =Yes 120 <conf:(1)> lift:(1) lev:(0) [0] conv:(0)
4. Do you use the internet for other things? =Yes 130 ==> Do you use the internet for your Academic Studies? =Yes 120 <conf:(0.92)> lift:(1.01) lev:(0.01) [0] conv:(1)
5. Do you use the internet for your Academic Studies? =Yes 130 ==> Do you use the internet for other things? =Yes 120 <conf:(0.92)> lift:(1.01) lev:(0.01) [0] conv:(1)
6. Do you use the internet for Social Networking (Facebook, Messenger, etc.)? =Yes Do you use the internet for other things? =Yes 130 ==> Do you use the internet for your Academic Studies? =Yes 120 <conf:(0.92)> lift:(1.01) lev:(0.01) [0] conv:(1)
7. Do you use the internet for your Academic Studies? =Yes Do you use the internet for Social Networking (Facebook, Messenger etc.)? =Yes 130 ==> Do you use the internet for other things? =Yes 120 <conf:(0.92)> lift:(1.01) lev:(0.01) [0] conv:(1)
8. Do you use the internet for other things? =Yes 130 ==> Do you use the internet for your Academic Studies? =Yes Do you use the internet for Social Networking (Facebook, Messenger, etc.)? =Yes 120 <conf:(0.92)> lift:(1.01) lev:(0.01) [0] conv:(1)
9. Do you use the internet for your Academic Studies? =Yes 130 ==> Do you use the internet for Social Networking (Facebook, Messenger, etc.)? =Yes Do you use the internet for other things? =Yes 120 <conf:(0.92)> lift:(1.01) lev:(0.01) [0] conv:(1)
10. Do you use the internet for Social Networking (Facebook, Messenger etc.)? =Yes 142 ==> Do you use the internet for your Academic Studies? =Yes 130 <conf:(0.92)> lift:(1) lev:(0) [0] conv:(0.92)

Table-2: Ranking among the factors with Ranking Factors

0.4591	12	Please indicate the extent to which you agree or disagree with each of the following statements. [b. The internet helped me achieved better academic performance]
0.4573	14	Please indicate the extent to which you agree or disagree with each of the following statements. [d. The internet is a reliable source for improving my academic performance]
0.3818	13	Please indicate the extent to which you agree or disagree with each of the following statements. [c. I get information relevant to my studies from the internet]
0.3255	11	Please indicate the extent to which you agree or disagree with each of the following statements. [a. I depend on the internet for most of my academic studies]
0.2511	19	The internet problems in my institutes
0.1843	2	Please indicate your education level
0.1742	20	When do you use the internet mostly?
0.156	8	How many hours a day do you use the internet For your Academic Studies?
0.1098	9	How many hours a day do you use the internet for your Social Networking?
0.1041	15	Please indicate the extent to which you agree or disagree with each of the following statements. [a. I usually spend more time on social networking sites than I should]
0.0905	10	How many hours a day do you use the internet for: (please circle one response only (e.g. Online gaming, Movie etc.)
0.0821	17	Please indicate the extent to which you agree or disagree with each of the following statements. [c. I cannot cut down the time I spend on social networking sites]
0.0791	4	Please indicate Your Age
0.0601	16	Please indicate the extent to which you agree or disagree with each of the following statements. [b. I go to sleep late at night because of the use of social networking sites]
0.0402	18	Please indicate the extent to which you agree or disagree with each of the following statements. [d. I get less time for academic studies due to the excessive use of social networking]
0.0384	5	Do you use the internet for your Academic Studies?
0.035	3	Please indicate Your Gender
0.0221	7	Do you use the internet for other things?
0	6	Do you use the internet for Social Networking (Facebook, Messenger, etc.)?

Table-2 shows the ranking from highest to lowest significance.

Accuracy of the dataset

Both Table 3 and Table 4 show the Confusion matrix and Accuracy tables of the dataset. Most of the results show better accuracy for Strongly Agreed.

Table 3: Accuracy Table

TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	Class
1.000	1.000	0.718	1.000	0.836	0.492	0.715	Strongly agree
0.000	0.000	0.000	0.237	0.035			Strongly disagree
0.000	0.000		0.000	0.452		0.216	Neutral
0.000	0.000	0.000		0.050		0.007	Strongly disagree
0.000	0.000	0.000	0.050	0.007			Neutral
Weighted Avg.	0.718	0.718	0.718		0.468	0.565	

Table 4: Confusion Matrix

a	b	c	d	e	<-- classified as
102	0	0	0	0	a = Strongly agree
5	0	0	0	0	b = Strongly disagree
33	0	0	0	0	c = Neutral
1	0	0	0	0	d = Strongly disagree;Strongly agree
1	0	0	0	0	e = Neutral;Strongly agree

Fig.-6 shows the Support Vector Machine analysis in which the red line indicates significance. The data pattern indicates that Internet usage has a higher impact on junior level male students compared to others.

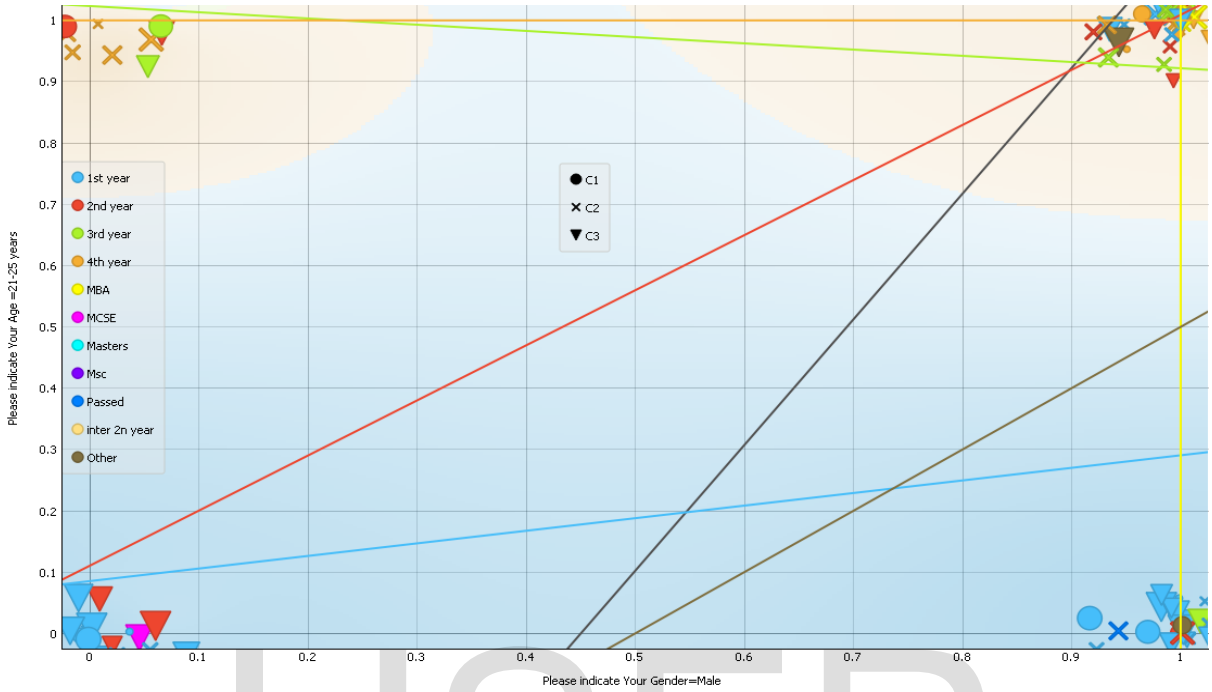

Fig.-6: Field pattern according to SVM analysis with different semester students

Decision Tree

A decision tree of the significant factors has been shown in Fig-7. The decision tree implies that Internet usage has a higher impact on the 2nd year students. From the tree, it is clear that the 1st year students use social networking media more than others.

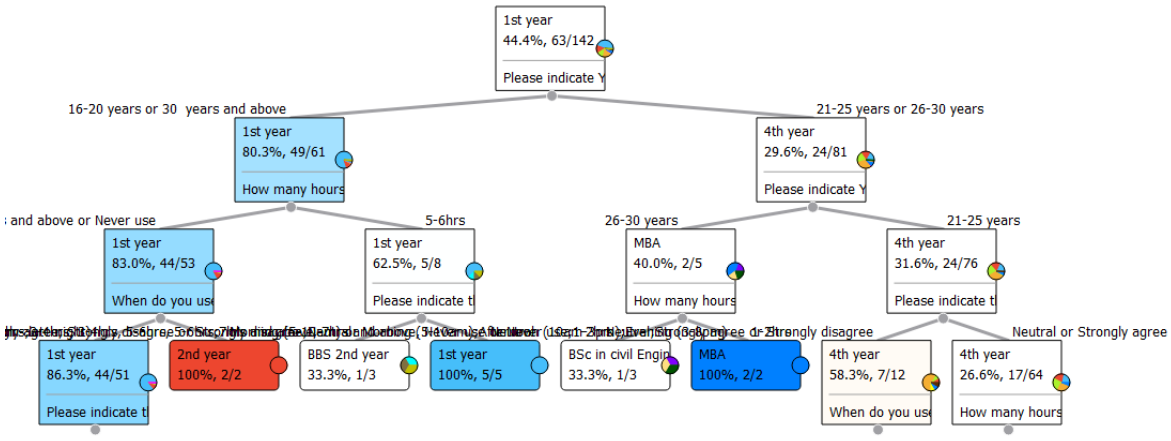

Fig.-7: Decision Tree of the significant factors

Addiction to Internet Usage

The important point was to see whether the subject students were addicted to Internet Usage and especially to networking sites. Fig.-8 shows that a simple majority (51%) of the responded students usually spend more time on social networking media, sleep late at night, and get less time for

academic purposes. About 38% of them have reported that they are addicted and cannot cut down on social networking media.

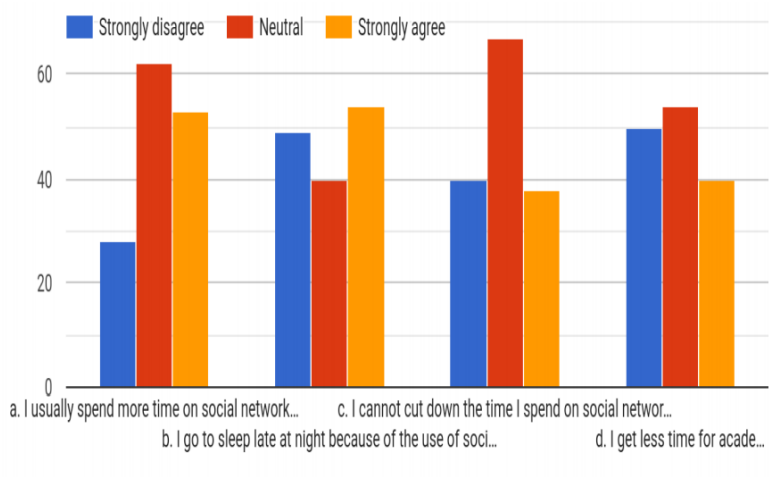

Fig. 8: Risky behaviors with Internet Usage

Impact on Academic Performance

Fig.9- shows the majority of the respondents (73%) have reported that Internet usage has improved their academic performance, while a small percentage about 5% of the respondents reported otherwise.

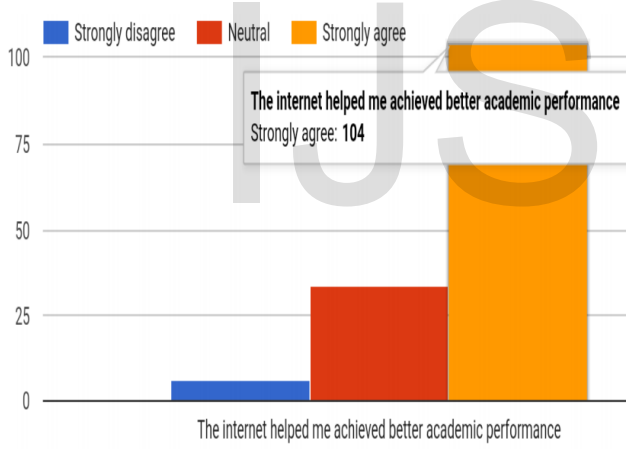

Fig. 9: Impact on Academic Performance

Conclusion

This research shows that overall results generally corroborate with the findings in which the participants use the Internet as useful tools in their academic studies (Emeka & Nyeche, 2016), and their academic performance improves(Muniandy, 2010).

The conclusion drawn from this research is that factors like social media use, internet problems, addiction to social networking, etc are affecting students and their academic performance. The impact of Internet usage on university students is high, and it impacts mostly junior male students.

Further study is suggested using a larger dataset for in-depth more accurate analysis and taking up some remedial programs for the students who are already adversely affected due to the addition of Internet usage.

Acknowledgment

The Authors are grateful to those who have participated with their data in this research.

Competing of Interest

Authors declare that they have no competing of interest among them

References

Akhter, N. (2013, October 10). Relationship between Internet Addiction and Academic Performance among University Undergraduates. *Academic Journals, 8*(19), 1793-1796. doi:10.5897/ERR2013.1539

Al-Jubayer, S. (2013, March). The Use of Social Networking Sites Among Teenagers: A Study of Facebook Use in Dhaka City. *Journal of International Social Issues, 2*(1), 35-44.

Assaduzzaman, S., & Al Masud, F. B. (2018). Dataset on significant risk factors for Type 1 Diabetes: A Bangladeshi perspective. 700-708.

Dimitriu, L. (2015). Changes in living cell morphology induced by electroporation. *Biointerface Rsearch in Applied Chemistry, 5*(5).

Emeka, U. J., & Nyeche, O. S. (2016, October). Impact of Internet Usage on the Academic Performance of Undergraduates Students: A case study of the University of Abuja, Nigeria. *International Journal of Scientific & Engineering Research, 7*(10).

Guan, S.-S., & Subrahmanyam, K. (2009, July). Youth Internet use: risks and opportunities. (R. W. HAZELL, Ed.) *CHILD AND ADOLESCENT PSYCHIATRY, 22*(4), 351-356. doi:10.1097/YCO.0b013e32832bd7e0

Hasebrink, U., Livingstone, S., Haddon, L., & Ólafsson, K. (2009). *Comparing children's online opportunities and risks across Europe: Cross-national comparisons for EU Kids Online.* London: LSE. Retrieved from http://ec.europa.eu/information_society/activities/sip/index_en.htm

Kader, R. (2019, January 31). The Mobile And Internet Penetration Growth Continues, Internet's Deployment Phase. *Future Startup*.

Krejcie, R., & Morgan, D. (1970). Determining Sample Size for Research Activities. *Educational and Psychological Measurement.*, 607-610.

Muniandy, B. (2010). Academic Use of Internet among Undergraduate Students: A Preliminary Case Study in a Malaysian University. *International journal of cyber society and education, 3*(2), 171 - 178. doi: 10.29430/IJCSE.201012.0171

Nugent, R., Shannon, R., McNamee, H., & Molyneaux, F. (2015). *A research study examining how young people's use of ICT and the amount of screentime impacts on GCSE attainment.* London: National Children's Bureau (NCB) : working with children, for children. Retrieved October 2015

Puspita, R. H., & Rohedi, D. (2018). The Impact of Internet Use for Students. *IOP Conf. Series: Materials Science and Engineering . 306.* North Sulawesi, Indonesia: IOP Publishing. doi::10.1088/1757-899X/306/1/012106

Senthil, V. (2018, June 6). Does the more internet usage provide good academic. *Education Information Technologies, 23*, 2901-2910. doi:https://doi.org/10.1007/s10639-018-9749-8

Siraj, H. H., Salam, A., Hasan, N. A., Jin, T. H., Roslan, R. B., & Othman, M. N. (2015, April). Internet Usage and Academic Performance: A Study in a Malaysian Public University. *International Medical Journal, 22*(2), 83 - 86.

Soegoto, E. S., & Tjokroadiponto, S. (2018). Effect of Internet on Student's Academic Performance and Social Life. *IOP Conf. Series: Materials Science and Engineering. 407.* IOP Publishing. doi:10.1088/1757-899X/407/1/012176

Stanislav, A. A., & Vassu-Dimov, T. (2009). Preliminary results of the investigations regarding the association of transforming growth factor-beta1 (tgfb1) gene polymorphism to metabolic syndrome in a romanian patients' group. *Biointerface Research in Applied Chemistry*(9), 3974-3978.

UNICEF. (2020). *One million school children to receive online safety certification by next year.* Dhakma: Press release by UNICEF.

Yilmaz, R., Yilmaz, F. K., Ozturk, H. T., & Karademir, T. (2017, February 6). Examining Secondary School Students' Safe Computer and Internet Usage Awareness: An Example from Bartın Province. *Pegem Eğitim ve Öğretim Dergisi,, 7*(1), 83-114. doi:10.14527/pegegog.2017.004