The Relationship between Students' Results and Spending Time on the Internet at Higher Education Level in Bangladesh

[1] Keya Rani Das, [2] Bishnu Kumar Adhikary, and [3] Khandoker Saif Uddin

[1] Department of Statistics, Bangabandhu Sheikh Mujibur Rahman Agricultural University, Gazipur-1706, Bangladesh, Email: keyadas57@bsmrau.edu.bd
[2] Institute of Education and Research (IER), University of Rajshahi, Rajshahi-6205, Bangladesh, Email: bkadhikary@ru.ac.bd
[3] Department of Statistics, International University of Business Agriculture and Technolgy, Uttara-1230, Bangladesh, Email: kh.saifuddin@hotmail.com

Abstract

Nowadays, the internet is a rapidly growing source of learning in education especially in higher education. Internet usage doesn't mean that the students of higher education use this only for academic purposes but use non-academic purposes also. A good number of researches revealed that there is a relationship between students' time spending on the internet and academic results. In this study, the researchers aim is to focus on the relationship between students' academic result and their internet usage times for different purposes. Data was collected by a close-ended questionnaire. Simple Random Sampling (SRS) is used to collect data from one hundred (100) undergraduate and graduate-level students. Data were collected from Bangabandhu Sheikh Mujibur Rahman Agricultural University, Gazipur, and the University of Rajshahi. Study variables are students' results, the time of spending on the internet uses for their academic and non-academic purpose. Correlation and regression analysis is widely applied for this educational data set. In this study, correlation and multiple regression model is applied in collected educational data with its assumption checking. Based on the value of R^2, *adjusted* R^2, *AIC, BIC,* and the standard error of the estimate a fitted model is presented in this study. All the coefficients in the model are statistically significant here and the F statistic shows the model is significant. From the fitted regression model, it has been found that the students' results changed positively when they spend time on the internet browsing for searching different types of educational and instructional materials for learning purposes. At the same time students' academic results degrading when they spend time browsing the internet for various non-academic activities.

Keywords— Students' Results, Academic, Non-academic, Internet use, Correlation, Assumption Checking, Higher Education.

Manuscript received on 23 January 2020 and accepted for publication on 4 November 2020.

1 INTRODUCTION

During the last decade, the academic infrastructure has become digital with increased interconnections among products, processes, and services. The internet now occupies a central role in any academic environment [1]. It is the doorways to explore the educational resources in higher education. It is a good opportunity to learn more besides the classrooms. Students are paying more time on the internet for their academic and non-academic resources from the past time [1]. They are getting more resources for their academic and non-academic purpose from the internet. The academic use of the internet is related to their curricula such as writing papers, searches for answers to questions, and preparing for assignments and others. On the other hand, non-academic activities mean playing games, showing videos, listening to music, upload and download recreational materials, etc. [2]. It can attract students' attention highly. Research shows that positive aspects of internet usage for communication, the internet being appreciated as an excellent medium of knowledge [3]. Studies mentioned that the internet has a direct influence on academic achievement with using educational software, and the provision of essential information [4, 5]. As a result, it is found that has a good relationship and impact on students' academic results because they learn very quickly and easily through the internet [6]. If we want to study this type of academic impact data, we need to validate it through scientific analyses. In this case, statistics is a must in every scientific research [7]. Its application in the scientific branch makes the significant test result as well as increases the research validity and precision. Without applying statistical verification, scientific research goes to lame and not precise [8]. In educational data analysis, different types of statistical methods are applied [9, 10]. Among these, correlation and regression analysis are widely used methods. Correlation and regression analysis is a very common technique to examine data sets and also make an inference about the population parameters involved in the model in the educational perspective. Students' results depend on some factors surely of academic activities. In this paper, the researchers aimed to focus on whether the relationship and impact between students' results and internet usage for their academic and non-academic purpose.

1.1 Research Questions

The main aim of this research is to find out the relationship between students' results and spending time on the internet on academic and non-academic activities in higher education in Bangladesh.

The specific research questions are:

a) Which will be a well-fitted model of students' results depend on their spending time on the internet for academic activities and non-academic activities?

b) What are the impacts of using the internet on academic and non-academic activities at the higher education level in Bangladesh?

2 Methodology

This study is sought to find out the relationship between students' results and internet uses for their academic and non-academic activities at the higher education level in Bangladesh. In this regard, correlation and regression analysis are applied here for the collected educational data set through SPSS and R. This is a quantitative study by nature. Simple Random sampling (SRS) was followed as the technique of data collection to fit with quantitative nature. One hundred (100) undergraduate and graduate-level students have selected randomly for data collection. Data were collected from Bangabandhu Sheikh Mujibur Rahman Agricultural University, Gazipur and Rajshahi University, Rajshahi of Bangladesh. A close-ended questionnaire was used as the method of data collection. Students' result (into Cumulative Grade Point Average-CGPA on the scale of 4.00) has considered as the depended variable. On the other hand, the time of spending on the internet use of students' for their educational purpose (in hours) and time of spending on the internet for various non-academic activities (in hours) are independent variables.

2.1 Correlation

Correlation and regression are very common terms in statistical analysis. For non-practitioners, sometimes it is hard to perform correlation and regression analysis and also to interpret. Correlation does not focus on cause-and-effect, but sometimes research is misguided on that issue [11]. We study correlation analysis to measure the strength of the linear relationship between two or more variables. By the coefficient of correlation symbolized by r, we measure the association or correlation among variables. Coefficient of correlation developed by Karl Pearson and generally referred to as *Pearson's r*. Firstly, we have to look at the scatter diagram to see whether the linear relationship exists or not [12]. The correlation coefficient ranges from -1 to1.

r can be defined as

$$r_{xy} = \frac{Cov(xy)}{\sqrt{Var(x)Var(y)}}$$

Basic idea about correlation and regression is also available in Zou et al. (2003); Seber and Lee (2003) [13, 14].

2.2 Regression

The term 'regression' is originated by famous Biometrician Sir F. Galton (1822-1911) with the inheritance of stature. Presently, regression is a very popular technique to apply in the educational research area. Correlation reveals the strength and direction of a linear relationship between two variables but regression express the nature of the relationship by a mathematical equation to estimate the value of dependent variable Y which is based on the selected value of independent variable X. There is a mathematical equation created by regression in where if one variable is known, then another variable can be estimated [15]. Necessary steps for regression analysis are available in Daniel (2012) [16].
Mathematically, a linear regression line creates an equation where we can estimate one variable if another variable is known to us.

The linear regression model is

$$Y_i = E(Y/X_i) + e_i = \alpha + \beta X_i + e_i; \; i = 1, 2, \ldots\ldots n$$

Where α = the intercept of the line on the Y-axis.

β = Regression coefficient or slope of the line indicating the change in the dependent variable for each unit change in the independent variable.

Y_i = Dependent variable.

X_i = Independent variable.

Here e_i's are independently normally distributed with the same variance.

In the linear regression analysis, we need to check some assumptions like normality, linearity, homoscedasticity, multicollinearity, and outlier.

a. Normality

It requires residual to be normal into the linear regression line. By histogram, Q-Q plot we can check the normality of the data set. Also, Kolmogorov-Smirnov test value explains that the data are normally distributed or non-normal. To inquire about the normality of residuals we can use a normal predicted probability plot (P-P) plot. If the data are non-normal it might be needed to apply the non-linear transformation. In the normal predicted probability plot, a diagonal line is there and some circle dots are surrounded by the diagonal line. It obeys the normality assumption if dots are following the normality line. In the same way, it does not obey the normality assumption if dots are deviating from the line more. The little deviation may consider, but extreme deviation from the line may not be considered for the normality assumption for the regression line. An excellent review is listed about normality and its test into some researches [17, 18, 19].

b. Linearity

In the regression line, it needs to be a straight line for linear regression analysis, which is termed as linearity of that model. A Scatter diagram helps to see the linearity of data. When residuals of the regression model obey normality assumption with a homoscedastic pattern at all, in that time we don't need to check this assumption as it follows a normal distribution and it will be linear.

c. Homoscedasticity

Homoscedasticity means that the variance on all sides of the regression line is identical for all values of the independent variable. If the residuals are uniformly distributed all over then it shows a homoscedastic pattern in the data. We can get an idea from a t-test or ANOVA about equality of variance. Also by the scatter diagram between residuals and predicted values we can check this assumption as equality of variance.

When the dot points are distributed uniformly above and below from the zero lines then it explains that the data are homoscedastic otherwise data are heteroscedastic. If the data does not follow this assumption which means the pattern of the data set is heteroscedastic then a non-linear data transformation might be needed for the data set.

d. Multicollinearity

For some data sets, the independent variables are highly correlated with each other which is termed as multicollinearity. The estimate of the standard error of regression coefficients increases and also imprecise estimates of regression coefficients with wrong signs due to multicollinearity problem [20]. It can be measured by *VIF* elaborated as the variance inflation factor. It suggests that if the value of *VIF* lies below 5 then there is no multicollinearity problem. Also, the correlation coefficient (r) value gives an idea of whether the independent variables are correlated with each other or not. After checking, this assumption researcher needs to discard the independent variables as per research requirements or they can be applied ridge regression to their data.

e. Outliers

Outliers are anomalous values that may play a potential influence over the fitted slope. Residual variance increases or makes a poor fit of the data points due to outliers. After fitting the regression line, the boxplot or normal probability plot for residuals can make suggestions about the presence of outliers in the data. If the data suffer from outliers then some researchers suggest removing those cases and some researchers suggest applying nonparametric or other alternative regression methods to the data. In the boxplot, data points outside the black line are considered to be extreme outliers and are marked with a different symbol [21].

2.3 Multiple linear regression analysis

Multiple regression predicts a single dependent continuous variable where a group of independent variables exists [22]. When independent variables are more than one, then we can apply multiple regression analysis. Sometimes multiple regression analysis is applied in the wrong way by educational researchers [23]. For two independent variables X_1 and X_2, the form of the multiple regression model is

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + e_i$$

Where β_i's are regression coefficients and Y is the Students' result. Also, X_1 is spending time on the internet for educational purposes and X_2 is spending time on the internet for non-academic purposes.

3 Result and Discussion

The following diagram shows a linear relationship between variables. Now we can apply correlation analysis for these variables.

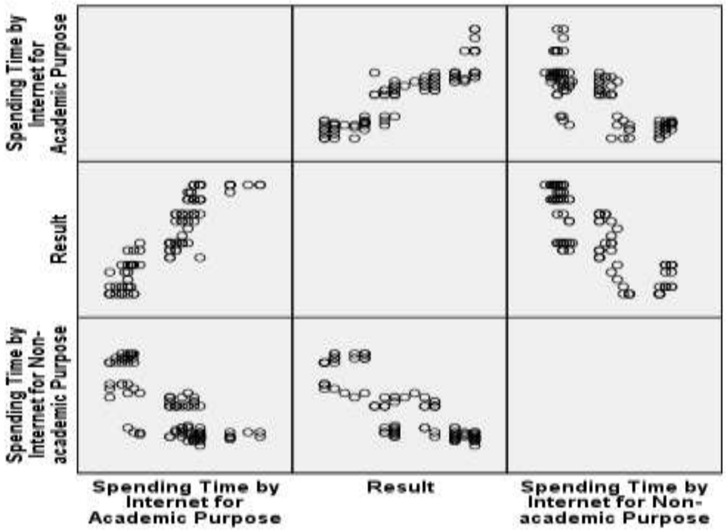

Figure 1. Matrix scatter plot

From this matrix scatter plot, there is a linear relationship between results and spending time on the internet for their educational purpose. Also, it shows an upward pattern which indicates a positive relationship between these variables. Again, it depicts a linear pattern between results and spending time on the internet for non-academic purposes but there is a downward pattern in points. So, it indicates a negative relationship between these two variables.

Table. 1: Correlations

		Spending Time by Internet for Academic Purpose	Result	Spending Time by Internet for Non-academic Purpose
Spending Time by Internet for Academic Purpose	Pearson Correlation	1	.91**	-.780**
	Sig. (2-tailed)		.000	.000
	N	100	100	100
Result	Pearson Correlation	.91**	1	-.823**
	Sig. (2-tailed)	.000		.000
	N	100	100	100
Spending Time by Internet for Non-academic Purpose	Pearson Correlation	-.780**	-.823**	1
	Sig. (2-tailed)	.000	.000	
	N	100	100	100

**. Correlation is significant at the 0.01 level (2-tailed).

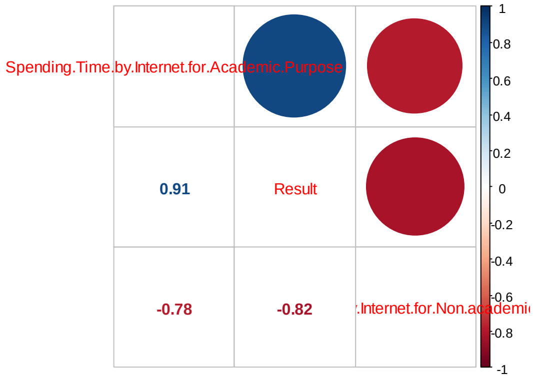

Figure 2. Correlation study

Here the Pearson correlation coefficient between Result (in CGPA) and Students' spending time on the internet for their academic purpose is .91 and the significance level is .000 which means it is highly significant. This *r* value tells that a strong positive correlation remains between these two variables. Again the Pearson correlation coefficient between Result (in CGPA) and Students' spending time on the internet for non-academic purposes is -0.823 and it is also highly significant. There is a strong negative correlation between these two variables.

A study has disclosed that there are only 19.43% and 33.18% of students can answer a good explanation about correlation and regression coefficient respectively [24]. Based on the theoretical concepts of this study, it suggested that correlation and regression analyses need to be understood more clearly by the students. In this regard, this article also featured a basic idea of correlation and regression analyses. A survey of 572 students from a public university was conducted which shows that excessive internet use is strongly correlated with low academic performances [25]. Again, compared with the normal uses of the internet it was found that heavy internet use has significantly lower academic results, social interactions, and learning satisfaction [26]. It was revealed that there was no significant difference in Internet Addiction (IA) in terms of cumulative Students' Academic Performance (SAP) [27]. This study was conducted among foreign undergraduate students in Malaysia. Digital models are transparent to all and students can learn quickly, because they get the opportunity to learn anything at any time on their hands [6]. It was also found that due to excessive internet usage students lose interest in academics, which leads to a decreased academic performance and a deterioration of the relationship with teachers [28]. Different researches showed that symptoms of internet addiction (IA) include preoccupation, withdrawal, loss of control, and functional impairment [29, 30].

Table. 2: Regression output

Subject	Result
Constant Value (*a*)	2.954 (0.000)
b_1 *VIF*	0.567 (0.000) 2.558
b_2 *VIF*	-0.222 (0.000) 2.558
R^2	0.854
Adjusted R^2	0.851
Standard error of the estimate	0.20
AIC	-32.86
BIC	-22.44
F statistic	283.34 (0.000)

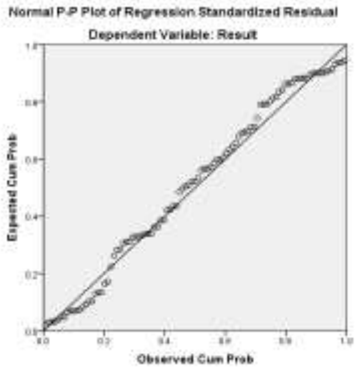

Figure 3. Normal P-P plot

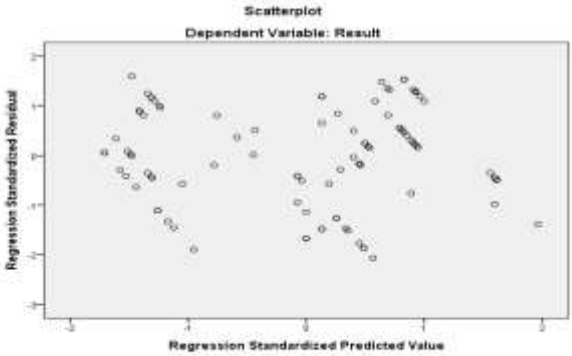

Figure 4. Scatter plot between residuals and predicted values

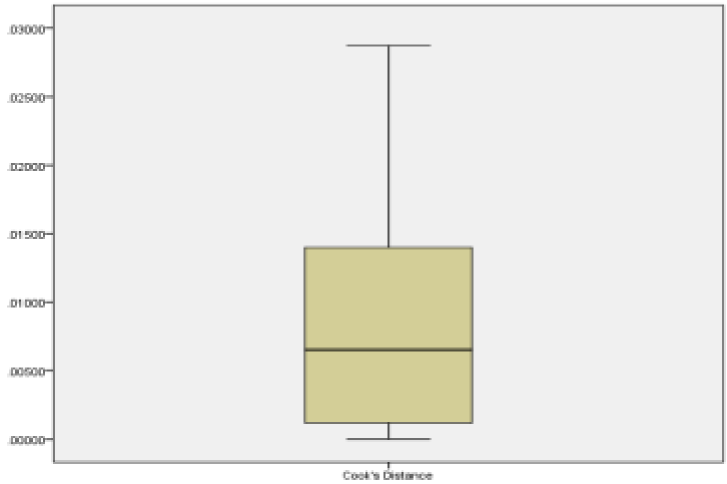

Figure 5. Boxplot of Cook's Distance

A prediction about the academic result on 150 students has been studied [31] and the linear regression model was applied to that study. The independent variables were spending time on the internet for different activities and the predicted model explained there. But in this research, we wanted to analyze students spending time on the internet for educational purposes and students' spending time on the internet for different non-academic purposes as our independent variables. After checking assumptions finally, we have been presented a fitted regression model here. This fitted regression model results (CGPA) = 2.954 + 0.567 (spending time by internet for academic purpose) – 0.222 (spending time by internet for non-academic purpose). The R^2 value is 0.854 and the adjusted R^2 value is 0.851 which are both present high values and it may be showed that the fitted regression model is good enough. Also, the standard error of the estimate for this model is 0.20 which is a very low value, and this low value also an indicator for a well-fitted model. *Akaike information criterion (AIC)* and *Bayesian Information Criterion (BIC)* are presented here for the regression model. All the coefficients as b_1 and b_2 are statistically significant here and also the F statistic shows the model is significant. Here multicollinearity is diagnosed by *VIF*. *VIF* for b_1 and b_2 are 2.558 which is less than 5. For this data, *VIF* depicts that multicollinearity does not arise here. Also, the normal p-p plot (in Figure 3) presents that it mostly follows the normality assumption. The scatter diagram of residuals and predicted values (in Figure 4) presents no pattern of the diagram for residuals and the coordinates are distributed uniformly above and below from the zero lines. For this data, a boxplot of *Cook's distance* (in Figure 5) shows that there is no star (*) mark outside of the black line. So this data possess no serious outlying problem. In this regard, the final fitted regression line depicts that students' results will be positively changed when they use internet browsing for searching different types of educational materials or simply academic purposes. Assumptions are checked properly for this data set and finally, the fitted regression model is presented in this study. So internet browsing for non-academic and academic purposes impacts on students' academic results negative and positive respectively.

4 Conclusion

Currently, the internet has moved into homes, schools, internet cafes, and businesses, there has been a rapidly growing public awareness of the potential adverse effects arising from excessive, maladaptive, or addictive Internet usage [32, 33]. In this study, the relationship on the internet usage of the higher education level students for academic and non-academic purposes are analyzed here. From our study result, based on students' time spending on the internet a fitted model has been identified. After checking important assumptions the fitted regression model explains students' result switching into good due to spending time on the internet browsing for academic learning purposes. The fitted regression model presents here it harms students' academic result when they browse the internet for different non-academic activities. There is a significant impact on students' academic results with internet usage for academic and non-academic purposes. A limitation of this study is confined to only two universities and a limited number sample size thus the generalization of findings to the entire population is limited. This limitation will be overcome in our future research work. This is an initial effort to study the relationship between students' time spending on the internet and their academic performances. It also offers noteworthy insights to the academicians and researchers.

References

[1] V. Senthil, "Does the more internet usage provide good academic grades?" Educ Inf Technol, vol. 23, pp. 2901–2910, 2018, https://doi.org/10.1007/s10639-018-9749-8.
[2] J. A. Moriana, A. Francisco, A. Rocio, P. Maria-Jose, H. Javier and R. Rosario, "Extra-Curricular activities and academic performance in secondary students," Electronic Journal of Research in Educational Psychology, vol. 4, no. 1, pp. 35-46, 2006.
[3] V. Stancis and A. Tinca, "A critical look on the student's internet use – An empirical study," Journal of Accounting and MIS, vol. 13, no. 4, pp. 739–754, 2014.
[4] O. Jegede, P. Okebukola and G. Aiewole, "Attitude to the use of the Computer for Learning Biological Concepts and Achievement of Students in an Dominated by Indigenous Technology," ERIC Document Reproduction Services, 1990.
[5] L. A. Jackson, A. von Eye, F. A. Biocca, G. Barbatsis, Y. Zhao and H. E. Fitzgerald, "Does home internet use influence the academic performance of low-income children?"Developmental Psychology, vol. 42, no. 3, pp. 429-435, 2006, DOI: 10.1037/0012-1649.42.3.429.
[6] P. Weill and S. L. Woernar, "Optimizing your digital business model,"MIT Slogan Management Review, vol. 54, no. 3, pp. 71–78, 2013.
[7] C. Weihs and K. Ickstadt, "Data Science: the impact of statistics,"Int J Data Sci Anal, vol. 6, pp. 189-194, 2018, DOI: https://doi.org/10.1007/s41060-018-0102-5.
[8] R. A. Oster and F. T. Enders, "The Importance of Statistical Competencies for Medical Research Learners," Journal of Statistics Education, vol. 26, no. 2, pp. 137-142, 2018, DOI: 10.1080/10691898.2018.1484674.
[9] S. Naskar and P. Das, "Applications of different Statistical Tests in Educational Research: An Overview," Journal of Emerging Technologies and Innovative Research, vol. 5, no. 5, pp. 129–137, 2018.
[10] J. W. Creswell, Educational research: Planning, conducting, and evaluating quantitative and qualitative research. 4th ed., Boston, MA: Pearson, 2012.
[11] L. Ludbrook, "Statistical Techniques For Comparing Measurers And Methods Of Measurement: A Critical Review, Clinical and Experimental Pharmacology and Physiology," Wiley Online, vol. 29, no. 7, pp. 527-536, 2002.
[12] V. Bewick, L. Cheek and J. Ball, "Statistics review 7: Correlation and regression," Critical Care, vol. 7, no. 6, pp. 451–459, 2003.
[13] K. H. Zou, K. Tuncali and S. G. Silverman, "Correlation and Simple Linear Regression," Radiology, vol. 227, no. 3, pp. 617-628, 2003.
[14] G. A. F. Seber and A. J. Lee, Linear regression analysis. 2nd ed., NY: Wiley, 2003.
[15] R. Winters, A. Winters and R. G. Amedee, "Statistics: A Brief Overview," The Ochsner Journal, vol. 10, no. 3, pp. 213–216, 2010.
[16] W. W. Daniel, BIOSTATISTICS: Basic Concepts and Methodology for the Health Sciences. 9th ed., Wiley India, pp. 413, 2012.
[17] K. R. Das and A. H. M. R. Imon, "A Brief Review of Tests for Normality," American Journal of Theoretical and Applied Statistics, vol. 5, no. 1, pp. 5-12, 2016, DOI:10.11648/j.ajtas.20160501.12.
[18] G. G. Judge, W. E. Griffith, R. C. Hill, H. Lutkepohl and T. Lee, Theory and Practice of Econometrics. 2nd ed., John Wiley & Sons, New York, 1985.
[19] F. Schmidt and C. Finan, "Linear regression and the normality assumption," Journal of Clinical Epidemiology, vol. 98, pp. 146-151, 2018.
[20] W. Yoo, R. Mayberry, S. Bae, K. Singh, Q. Peter He and J. W. Lillard Jr, "A Study of Effects of MultiCollinearity in the Multivariable Analysis," International Journal of applied science and technology, vol. 4, no. 5, pp. 9–19, 2014.
[21] R. Dawson, "How Significant Is A Boxplot Outlier?" Journal of Statistics Education, vol. 19, pp. 1-12, 2011.
[22] J. Pallant, SPSS survival manual, a step by step guide to data analysis using SPSS for windows. 3rd ed., McGraw Hill Sydney, pp 146, 2007.
[23] S. C. Karpen, "Misuses of Regression and ANCOVA in Educational Research," American Journal of Pharmaceutical Education, vol. 81, no. 8, pp. 6501, 2017, DOI: 10.5688/ajpe6501.
[24] F. A. Razak, N. Baharun, N. A. Deraman and N. R. P. Ismail, "Assessing Students' Abilities in Interpreting the Correlation and Regression Analysis," Journal of Fundamental and Applied Sciences, vol. 9, no. 5S, pp. 644-661, 2017, DOI: http://dx.doi.org/10.4314/jfas.v9i5s.45.
[25] R. W. Kubey, M. J. Lavin and J. R. Barrows, "Internet usage and collegiate academic performance decrements: Early findings," Journal of Communications, vol. 51, no. 2, pp. 366–382, 2011.

[26] Y. F. Chen and S. S. Peng, "University students' internet use and its relationships with academic performance, interpersonal relationships, psychosocial adjustment, and self-evaluation," Journal of Cyber psychological Behaviour, vol. 11, no. 4, pp. 467–469, 2008, https://doi.org/10.1089/cpb.2007.0128.

[27] N. H. Usman, M. Alavi, and S. M. Shafeq, "Relationship between internet addiction and academic performance among foreign undergraduate students," Procedia-Social and Behavioral 114, 4th World Conference on Psychology, Counseling and Guidance WCPCG-2013, pp. 845–851, 2014.

[28] O. A. Bankole, M. Lalitha, H. U. Khan, and Jingo, "Information technology in the maritime industry past, present and future: Focus on long carriers," 7th IEEE International Advance Computing Conference, Hyderabad, India, 2017.

[29] R. Tao, Huang, J. Wang, H. Zhang, Y. Zhang and M. Li, "Proposed diagnostic criteria for internet addiction," Addiction, vol. 105, pp. 556–564, 2010.

[30] O. Turel, A. Serenko and P. Giles, "Integrating technology addiction and use an empirical investigation of online auction users," MIS Quarterly, vol. 35, no. 4, pp. 1043–1061, 2011.

[31] R. R. Rajalaxmi, P. Natesan, N. Krishnamoorthy and S. Ponni, "Regression Model for Predicting Engineering Students Academic Performance," International Journal of Recent Technology and Engineering (IJRTE), vol. 7, no. 6S3, pp. 71-75, 2019.

[32] J. J. Block, "Issues for DSM-V: internet addiction," Am J Psychiatry, vol. 165, pp. 306-307, 2008.

[33] G. M. Cooney and J. Morris, "Time to start taking an internet history?" British Journal of Psychiatry, vol. 194, no. 2, pp. 185-185, 2009, https://doi.org/10.1192/bjp.194.2.185.

Keya Rani Das is an Assistant Professor of Department of Statistics, Bangabandhu Sheikh Mujibur Rahman Agricultural University (BSMRAU) at Gazipur. She has completed her B.Sc major in Statistics and M.Sc in Statistics from the University of Rajshahi, Bangladesh. She has eight national and international publications with high indexing research articles to her credit. Her research interests are Multivariate Analysis, Educational Statistics, Statistical Quality Control, and Bioinformatics. She is a Life Member of the Bangladesh Association for the Advancement of Science, Bangladesh.

Bishnu Kumar Adhikary is an Assistant Professor at the Institute of Education and Research (IER) of the University of Rajshahi (RU). Before joining IER, University of Rajshahi, he worked as a Researcher at the Institute of Education and Research (IER), University of Dhaka (DU). Adhikary did his M.Phil in 2016, Master of Education (M.Ed) and Bachelor of Education (Honours) the IER, University of Dhaka, Bangladesh. Recently he had completed the Inspire Round 4 international strategic partnership project of the British Council in research and education with the Open University, UK and IER, RU, University of Rajshahi. His research articles and presented papers mainly focus on education policy and development, social inclusion in education, language education, teacher education, and indigenous education. He has some research articles and Co-Author of a Teachers' Curriculum Guide (TCG), National Curriculum and Textbook Board.

Khandoker Saif Uddin is the Professor and Head of Department of Statistics, International University of Business Agriculture and Technology (IUBAT), Bangladesh. Currently, he is the Director of IQAC of this university. Before joining here, he was the Professor and Head of Department of Statistics, Bangabandhu Sheikh Mujibur Rahman Agricultural University (BSMRAU), Gazipur. He has seventeenth publications to his credit.