Enhanced Knowledge in Sciences and Technology Vol. 1 No. 1 (2021) 52-60

EKST

Homepage: http://publisher.uthm.edu.my/periodicals/index.php/ekst
e-ISSN : 2773-6385

Modelling Binary Logistic Regression For Time Management On The Students' Academic Performance

Norziha Che Him[1*], Nur 'Aliza Azmi[2], Suliadi Firdaus Sufahani[3], Yusliandy Yusof[4],

[1*,2,3]Department of Mathematics and Statistics,
Faculty of Applied Sciences and Technology,
Universiti Tun Hussein Onn Malaysia,
84600 Pagoh, Johor, MALAYSIA

[4]Kolej Kemahiran Tinggi MARA,
Batu 11, Jalan Kluang-Batu Pahat,
86400 Parit Raja, Johor, MALAYSIA

*Corresponding Author : norziha@uthm.edu.my

DOI: https://doi.org/10.30880/ekst.2021.01.01.008
Received 22 January 2021; Accepted 30 January 2021; Available online 11 February 2021

Abstract: Academic performance of students can be measured by using Cumulative Grade Point Average (CGPA) to look on how well the performance of the students in accomplish their continuous assessment in their programme. Overall, this study's purpose is to determine the effect of time management on the students' academic performance in Universiti Tun Hussein Onn Malaysia (UTHM), Pagoh Campus. A total of 357 questionnaires have been distributed to students in UTHM Pagoh Campus by using stratified random sampling. Response variable is the CGPA and the explanatory variables are the Time Management. Result shows that three most activities that involved by student known as academic, social and co-curriculum, meanwhile, three least activities by student are part-time, co-curriculum and social. Chi-Square Test of Independence presents four explanatory variables with p-value > 0.05 which are Having Daily Routine, Emphasise Academic-Related Activities, Planning Time To Relax and Shopping Or Sightseeing At Least Once A Week that indicated the null hypothesis accepted with the conclusion there is no significant association between explanatory variables to the response variable. Meanwhile, the explanatory variables such as Gender, Age, Races, Year of Study, Most Activity, Least Activity, Having Trouble In Organising Things, Finish The Assignments Within The Time Given, Use Time Optimally And Prioritize Important Task, Work Part Time On Weekends and Exercise Everyday recorded the p-value < 0.05 as indication of there is a significant association with the response variable. Finally, two logistics regression models known as Model A and Model B have been developed, but only Model A has been chosen as a best model with the lowest value of Akaike Information Criterion (AIC).

Corresponding author: norziha@uthm.edu.my

Keywords: Time Management, Academic Performance, Binary Logistic Regression, Cumulative Grade Point Average, Akaike Information Criterion

1. Introduction

There is no doubt that education still one of the most important aspect for human life and student plays main role as a future leader for the nation development [1]. Student academic performance can be measured by using Cumulative Grade Point Average (CGPA) and compute by dividing the summation of credit units' time's grade points with total number of modules taken by students. CGPA could be used to look at on how well the performance of students in accomplished their continuous assessment of whole programme. Overall, time management is a serious aspect that leave great impact toward nation education. Each person requires self learning on how to manage their time since it could be a good practice to help everyone in improving the quality of life. Time management could include arranging, organising, scheduling and budgeting time with the purpose to generate and accomplish any task to be more effective and productive. By focusing on the time management, it could help one to solve any future problem such as unable to deal with distractions, procrastination, deadline pressure and any others problems that related to time management [2].

Time management expected to have various effect to academic performance where it increase when students able to manage and well handle their time management. Expert believed it's important for students to understand on how to make a good decision in managing their time since academic performance may affected without proper schedule or planning [3]. [4] also agreed that the time management plays vital role in improving students' academic performance where majority of students that reported to have problem in their time management in college since their will obtain many responsibility towards their courses and programme if compared to the high school level. Research by [5] shows that students' academic success affected by effective time management and this supported by [6] where the time management has a positive influenced on academic performance and become main key to obtain positive result in college. This statement is also supported by [14] and [7] which they agreed that an effective time management become important in students' life since students' CGPA will increase if they supported by excellent time management skills.

This study aims to identify the types of activities and time spend for the activity, to investigate the relationship between academic performance of students and their time management and to develop binary logistic regression model on time management towards students' academic performance.

2. Materials and methods

2.1 Materials

A total of 357 questionnaires have been distributed to students in UTHM Pagoh Campus by using stratified random sampling. The questionnaire consists of two parts known as Part A (5 Questions) and Part B (15 Questions). The reliability of the questionnaire has been measured by using Cronbach's alpha with value 0.74 indicates that the questionnaire is reliable. Response variable is the CGPA and the explanatory variables are the time management.

2.2 Methods

According to [12], contingency table or cross tabulation of the observed frequencies of the two variables could estimate the association between two variables before continued with analysis technique by using Chi-Square Test of Independence. The following are the hypothesis testing of this study [11].

H_0 : There is no association between the two variables
H_1 : There is an association between the two variables

Multicollinearity is a situation where the predictor variables tend to have high correlation between them [10] and [13] highlighted that predictors could have Variance Inflation Factor (VIF) value more than 10, need to be removed since it indicates that there is very high correlation among variables. Meanwhile, Generalised Linear Model (GLM) extend the linear modelling framework to variables that are not Normally distributed. GLM are most commonly used to model binary or count data and GLM is made up of a linear predictor, $\eta_i = \beta_0 + \beta_1 X_{1i} + \ldots + \beta_p X_{pi}$ and two functions, which is a link function that describes how the mean, $E(Y_i) = \mu_i$, that depends on the linear predictor, $g(\mu_i) = \eta_i$ and a variance function that described how the variance, $\text{var}(Y_i)$ depends on the mean, $\text{var}(Y_i) = \phi V(\mu)$, where the dispersion parameter ϕ is a constant [17].

2.3 Equations

Logistic regression is one of the popular methods to analyse binary response variable other than the probit model and discriminant analysis [8]. In addition, Logistic regression is used to obtain odds ratio in the presence of more than one explanatory variable where the procedure is quite similar to multiple linear regression, except the response variable is binomial. The result is the impact of each variable on the odds ratio of the observed event of interest. The main advantage is to avoid confounding effects by analysing the association of all variables together [18]. Meanwhile, a binary logistic regression was applied when the response variable is dichotomous while the independent variables are either continuous or categorical. Several link functions are available, but the most common link for binary is the logit link function, defined as in Eq. 1.

$$g(\mu_i) = \text{logit}(\mu_i) = \log\left(\frac{\mu_i}{1-\mu_i}\right) \qquad Eq.\ 1$$

Thus, the model equation in logit form is given as in Eq. 2 [9],

$$\log\left(\frac{\mu_i}{1-\mu_i}\right) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p \quad Eq.\ 2$$

where, α is an intercept parameter, β is a regression coefficient, μ_i is the probability of an event occurring and $1-\mu_i$ is the probability of an event not occurring. Akaike Information Criterion (AIC) is a function of its maximized log-likelihood and the number of estimable parameters (K) and it often used to compare two or more GLM model [16]. The model with lowest value of AIC will select as the best model compared to model with highest value of AIC. Eq. 3 shows the formula of AIC.

$$AIC = 2\log-\text{likelihood} + 2K \qquad Eq.\ 3$$

Hosmer-Lemeshow test is commonly used to test the goodness-of-fit in the logistic regression where the test indicates the appropriateness of the model when the *p*-value is greater than 0.05 shows no evidence of poor fit model [16].

3. Result and Discussion

3.1 Descriptive statistics

This study focusing on the effect of time management towards students' academic performance. Majority of the students in UTHM Pagoh Campus performed well in their study and 243 students out of 357 students successfully obtained Cumulative Grade Point Average (CGPA) higher than 3.00. Here, CGPA is a response variable where students with CGPA lower than 3.00 been coded as 0, while coded

as 1, regards as students that have CGPA higher than 3.00. Figure 1 (left) shows three most activities that involved by student known as academic, social and co-curriculum. Meanwhile, Figure 1 (right) obviously presents that the three least activities by student are part-time, co-curriculum and social.

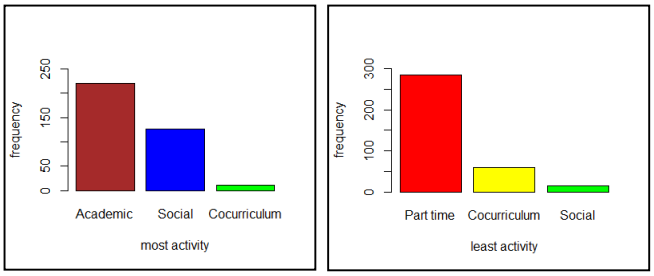

Figure 1: Distribution of respondents based on most activity (left) and least activity (right)

3.2 Association between time management and students' academic performance

The Chi-Square test has been adopted to determine the association between the time management and students' academic performance where the null hypothesis assumed that there is no significant relationship between response and explanatory variables, meanwhile for the alternative hypothesis assumed that there is a significant relationship between response and explanatory variables. Table 1 presents the Chi-Square Test of Independence with related explanatory variables and result. Obviously, there are four explanatory variables with p-value > 0.05 which are Having Daily Routine (β_{11}), Emphasise Academic-Related Activities (β_{13}), Planning Time To Relax (β_{15}) and Shopping Or Sightseeing At Least Once A Week (β_{18}) that indicated the null hypothesis accepted. Therefore, conclusion can be made that there is no significant association between explanatory variables to the response variable. Meanwhile, the explanatory variables such as Gender (β_1), Age (β_2), Races (β_3), Year of Study (β_4), Most Activity (β_5), Least Activity (β_7), Having Trouble In Organising Things (β_{10}), Finish The Assignments Within The Time Given (β_{12}), Use Time Optimally And Prioritize Important Task (β_{14}), Work Part Time On Weekends (β_{16}) and Exercise Everyday (β_{17}) recorded the p-value < 0.05 as indication of there is a significant association with the response variable.

Table 1: Chi-Square test of independence

Explanatory Variables	Coefficient	χ^2	p-value
Gender	β_1	4.0489	< 0.05
Age	β_2	8.8298	< 0.05
Races	β_3	62.9300	< 0.05
Year Of Study	β_4	9.7505	< 0.05
Most Activity	β_5	165.03	< 0.05
Least Activity	β_7	28.304	< 0.05
Having Trouble In Organising Things	β_{10}	41.1700	< 0.05
Having Daily Routine	β_{11}	2.4265	**> 0.05**
Finish The Assignments Within The Time Given	β_{12}	8.6607	< 0.05
Emphasise Academic-Related Activities	β_{13}	3.4476	**> 0.05**
Use Time Optimally And Prioritize Important Task	β_{14}	6.0263	< 0.05
Planning Time To Relax	β_{15}	2.1305	**> 0.05**

Work Part Time On Weekends	β_{16}	19.9890	< 0.05
Exercise Everyday	β_{17}	10.8820	< 0.05
Shopping Or Sightseeing At Least Once A Week	β_{18}	3.4218	**> 0.05**

3.3 Model development and comparison

Table 2 and Table 3 show the final selected explanatory variables and their estimate coefficient, standard errors and *p*-value for Model A with the equation as in Eq. 4. Therefore, we can see that variables No Technique Used ($\beta_{9(1)}$), More Than 5 Hours For Most Activity, (β_6), Age Group 21-23 ($\beta_{2(1)}$), Exercises Everyday (β_{17}) and Finish The Assignments Within The Time Given (β_{12}) have a negative relationship to CGPA. Meanwhile, a positive relationship towards CGPA involved following variables such as Most Activity (Academic) ($\beta_{5(0)}$), Most Activity (Social) ($\beta_{5(1)}$), Age Group 18-20 ($\beta_{2(0)}$), Use Time Optimally And Prioritize Important Task (β_{14}), Malay Race ($\beta_{3(0)}$), Chinese Race ($\beta_{3(1)}$) and Indian Race ($\beta_{3(2)}$).

Table 2: Final selected explanatory variables for Model A

Coefficient	Variables
α	Intercept
$\beta_{5(0)}$	Most Activity (Academic)
$\beta_{5(1)}$	Most Activity (Social)
$\beta_{9(1)}$	No Technique Used
β_6	More Than 5 Hours For Most Activity
$\beta_{2(0)}$	Age Group For 18-20
$\beta_{2(1)}$	Age Group For 21-23
β_{17}	Exercises Every Day
β_{14}	Use Time Optimally And Prioritize Important Task
$\beta_{3(0)}$	Malay Race
$\beta_{3(1)}$	Chinese Race
$\beta_{3(2)}$	Indian Race
β_{12}	Finish The Assignments Within The Time Given

Table 3: Estimates of coefficients, standard errors and *p*-values for Model A

Variables	Coefficient	Estimate	Std. error	Z-value	*Pr* (>Z)
Intercept	α	-0.101	1.626	-0.062	< 0.1
Most Activity (Academic)	$\beta_{5(0)}$	3.236	0.411	7.857	< 0.001
Most Activity (Social)	$\beta_{5(1)}$	1.316	0.803	1.638	< 1.0
No Technique Used	$\beta_{9(1)}$	-1.465	0.402	-3.644	< 0.01
More Than 5 Hours For Most Activity	β_6	-1.108	0.372	-2.978	< 0.01
Age Group For 18-20	$\beta_{2(0)}$	1.827	1.384	1.320	< 0.05

Age Group For 21-23	$\beta_{2(1)}$	-0.687	0.366	-1.877	**> 0.05**
Exercises Every Day	β_{17}	-0.590	0.245	-2.406	< 0.05
Use Time Optimally And Prioritize Important Task	β_{14}	0.670	0.302	2.214	< 0.05
Malay Race	$\beta_{3(0)}$	3.083	1.308	2.357	< 0.05
Chinese Race	$\beta_{3(1)}$	2.722	1.272	2.139	< 0.05
Indian Race	$\beta_{3(2)}$	1.619	1.214	1.333	**> 0.05**
Finish The Assignments Within The Time Given	β_{12}	-0.653	0.324	-2.013	< 0.05

$$\log\left(\frac{p}{1-p}\right) = -0.1010 + 3.24\beta_{5(0)} + 1.32\beta_{5(1)} - 1.47\beta_{9(1)} + 1.11\beta_{6} + 1.83\beta_{2(0)} - 0.69\beta_{2(1)} - 0.59\beta_{17} + 0.67\beta_{14} + 3.08\beta_{3(0)} + 2.72\beta_{3(1)} + 1.62\beta_{3(2)} - 0.65\beta_{12} \quad Eq.\ 4$$

Table 4: Final selected variables for Model B

Coefficients	Covariates
α	Intercept
β_{10}	Having Trouble In Organising Things
β_{16}	Work Part Time On Weekends
β_{19}	Satisfied In Managing Time
β_{17}	Exercise Everyday
β_{14}	Use Time Optimally And Prioritize Important Task

Table 5: Estimates of coefficients, standard errors and *p*-values for Model B

Variables	Coefficients	Estimate	Std. error	Z-value	*Pr* (>Z)
Intercept	α	1.203	0.836	1.439	> 0.05
Having Trouble In Organising Things	β_{10}	-0.584	0.141	-4.131	< 0.1
Work Part Time On Weekends	β_{16}	-0.323	0.177	1.772	< 0.05
Satisfied In Managing Time	β_{19}	0.400	0.133	-2.405	< 0.05
Exercise Everyday	β_{17}	-0.316	0.138	-2.289	< 0.05
Use Time Optimally And Prioritize Important Task	β_{14}	0.315	0.172	2.235	< 0.001

$$\log\left(\frac{p}{1-p}\right) = 1.203 - 0.585\beta_{10} + 0.315\beta_{14} - 0.323\beta_{16} - 0.317\beta_{17} + 0.400\beta_{19} \quad Eq.\ 5$$

In addition, Table 4 and Table 5 present the final selected explanatory variables and their estimate coefficient, standard error and *p*-value for Model B with the equation as in Eq. 5. Here, we can conclude that Having Trouble In Organising Things (β_{10}), Work Part Time On Weekends (β_{16}) And Exercise Everyday (β_{17}) Have A Negative Relationship With CGPA. Meanwhile, It Obviously Show That The Rest Two Variables, Satisfied In Managing Time (β_{19}) and Use Time Optimally And Prioritize Important Task (β_{14}) have a positive relationship with the response variable CGPA.

Table 6: Deviance, number of parameter (p), degrees of freedom (n-p) and Akaike Information Criterion (AIC) for Model A and Model B

Model	Deviance	*p*	*n-p*	AIC
Model A	445.7	19	338	249.70
Model B	445.7	10	347	412.17

A comparison model for Model A and Model B has been shown in Table 6. Obviously, Model A presents lowest AIC value that indicates a good model compared to Model B. This supported by Table 7 that clearly shows there is no evidence of poor fit for both models since the *p*-value > 0.05 for Hosmer-Lemeshow which are 0.129 and 0.582 respectively.

Table 7: The *p*-value for Hosmer-Lemeshow of Model A and Model B

Model	DF	Chi-Square	*p*-value
Model A	8	12.52	0.129
Model B	8	6.59	0.582

4. Conclusion

This study present majority of the students in UTHM Pagoh Campus spend more than five hours in activity that related to academic and two hours for part time working. The Chi-Square Test of Independence with related explanatory variables shows four explanatory variables with *p*-value > 0.05 which are Having Daily Routine (β_{11}), Emphasise Academic-Related Activities (β_{13}), Planning Time To Relax (β_{15}) and Shopping Or Sightseeing At Least Once A Week (β_{18}) that indicated the null hypothesis accepted, where there is no association to CGPA. The explanatory selected for Model A were Most Activity (Academic) ($\beta_{5(0)}$), Most Activity (Social) ($\beta_{5(1)}$), No Technique Used ($\beta_{9(1)}$), More Than 5 Hours For Most Activity (β_6), Age Group For 18-20 ($\beta_{2(0)}$), Age Group For 21-23 ($\beta_{2(1)}$), Exercises Every Day (β_{17}), Use Time Optimally and Prioritize Important Task (β_{14}), Malay race ($\beta_{3(0)}$), Chinese race ($\beta_{3(1)}$), Indian race ($\beta_{3(2)}$) and Finish The Assignments Within The Time Given (β_{12}). Meanwhile, for Model B, there were five explanatory variables that have been selected including Having Trouble In Organising Things (β_{10}), Work Part Time On Weekends (β_{16}), Satisfied In Managing Time (β_{19}), Exercise Every Day (β_{17}) And Use Time Optimally and Prioritize Important Task (β_{14}). Model A and Model B have been successfully developed, however only Model A was found to be the best model since the lowest AIC recorded value compared to Model B which 249.7 and 412.17 respectively.

Acknowledgement

This study would like to gratefully acknowledge Faculty of Applied Sciences and Technology (FAST) and Universiti Tun Hussein Onn Malaysia (UTHM) for moral support to publish this study and all respondent who successfully fill-in the survey form.

References

[1] A. M. Dahie and R. A. Mohamed, “Time Management And Academic Performance: Empirical Survey From High Education In Mogadishu-Somalia,” International Journal in Management and Social Science, November, pp. 376-388, 2016.

[2] F. Aduke, “Time Management And Students Academic Performance In Higher Institutions, Nigeria - A Case Study Of Ekiti State,” International Research in Eduaction, vol. 3, no. 2, pp. 1-12, 2015.

[3] P. M. Kaushar, “Study Of Impact Of Time Management On Academic Performance Of College Students,” Journal of Business and Management, vol. 9, no. 6, pp. 59–60, 2013.

[4] W. Siddiq, Y. Alsalem, L. A. Alamodi, A. Taher and M. Hazazi, “The Effect Of Time Management On Academic Performance Among Students Of Jazan University,” The Egyptian Journal of Hospital Medicine, vol. 69, no. October, pp. 3042–3049, 2017.

[5] A. Tesser, B. K. Britton and A. Tesser, “Effects Of Time-Management Practices On College Grades Effects Of Time-Management Practices On College Grades,” Journal of Educational Psychology, vol. 83, no. 3, pp. 405-410, 2014.

[6] K. Babayi, N. Hajloo, N. Sobhi and H. Sadeghi, “The Study Efficacy Of Time Management Training On Increase Academic Time Management Of Students,” Procedia - Social Behavioral Sciences, vol. 84, pp. 134–138, 2013.

[7] M. St, T. H. Macan, R. L. Dipboye and A. P. Phillips, “College Students Â€TM Time Management : Correlations With Academic Performance And Stress,” Journal of Education Psychology, vol. 82, no. 4, pp. 760–768, 1990.

[8] S. M. Hailpern and P. F. Visintainer, “Odds Ratios And Logistic Regression: Further Examples Of Their Use And Interpretation,” Stata Journal, vol. 3, no. 3, pp. 213–225, 2003.

[9] B. O. Sule, “A Logistic Regression Model Of Students’ Academic Performance In University Of Maiduguri, Maiduguri, Nigeria,” Mathematical Theory and Modelling, vol. 5, no. 10, pp. 124–137, 2015.

[10] R. K. Paul, "Multicollinearity: Causes, Effects And Remedies," Roll No. 4405, pp. 1-14, 2005.

[11] A. Kosťa´lova´, "Application Of Chi-Square Test Of Independence In The Utilization Of Postal And Telecommunication Services," The 10th International Conference Reliability and Statistics in Transportation and Communication, pp. 163-171, 2010.

[12] S. Amiri and R. Modarres, "Comparison Of Tests Of Contingency Tables," Journal of Biopharmaceutical Statistics, pp. 1-39, 2017.

[13] L. Murrey, H. Nguyen, Y. Lee, M.D. Remmenga and D.W. Smith, "Variance Inflation Factors In Regression Models With Dummy Variables," Conference on Applied Statistics in Agriculture, pp. 161-177, 2012.

[14] E.S. Indreica, A.M. Cazan and C. Truta, "Effects Of Learning Styles And Time Management On Academic Achievement," Procedia-Social and Behavioral Sciences 30, pp. 1096-1102, 2011.

[15] P. Paul, M. L. Pennell and S. Lemeshow, "Standardizing The Power Of The Hosmer-Lemeshow Goodness Of Fit Test In Large Data Sets," Statist. Med, 1-14, 2012.

[16] D. Posada and T.R. Buckley, "Model Selection And Model Averaging In Phylogenetics: Advantages Of Akaike Information Criterion And Bayesian Approaches Over Likelihood Ratio Tests," Systematic Biology, vol. 53, no. 5, pp. 793-808, 2004.

[17] A. J. Dobson, "An introduction to generalized linear models, 2nd ed.," Chapman & Hall/CRC texts in statistical science, 1945.

[18] S. Sperandei, "Understanding logistic regression analysis," Biochemia medica, vol. 24, no. 1, pp. 12–18, 2014.