Predicting Student’s Re-enrollment in Open and Distance Learning Environment

Sugilar

Universitas Terbuka

The Author Note

Sugilar, Department of Mathematics Education, Universitas Terbuka, is currently head of Bengkulu Regional Services of Universitas Terbuka. Correspondence concerning this article should be addressed to gilar@ut.ac.id

Predicting Student’s Re-Enrollment Odds in Open and Distance Learning Environment

Abstract

Student in open and distance learning environment has greater flexibility to re-enroll or not in every semester that make the students’ re-enrollment pattern is attractive to be studied.The purpose of this study was to predict student of Universitas Terbuka in one semester to re-enroll in the next semester from predictor variables consisting of student’s participating in learning support services, academic achievement, and personal characteristics. This research was conducted in quantitative approach. The sample was 1195 students taken randomly from the students of Universitas Terbuka who enrolled in the 1^{st}semester of 2015. Binary logistic regressionanalysis was employed to predict the probability that a new student in the 1^{st} semester of 2015 would enroll in the 2^{nd} semester of 2015. A test of the full model versus a model with intercept only was statistically significant, χ^2=275.470(df = 9, n = 1195), p <0.01). The model was classifying correctly 96.5% of students who enrolled and 40.8% of those who did not, for an overall success rate of 88.6%. It was finally found that (1) student participating in face-to-face tutorial had four more times likely to re-enroll, (2) student earning more than 12 credits had four times more chance to re-enroll, and (3) student achieved GPA more than 2.5 had almost three times more likely to re-enroll.

Keywords: Open Distance Learning, Universitas Terbuka, Prediction, Re-enrollment, Binary logistic regression.

Predicting Student’s Re-Enrollment Odds in Open and Distance Learning Environment

Open distance learning (ODL) has a long history for providing access for people who cannot participate in conventional higher education. However, Sharma (2015) quoted the president of International Council of Distance Education (ICDE), Tian Belawati, that the twin messages of open distance learning are access and success. Access is related to easiness and flexibility for people to participate in higher education, on the other hand, success,in most cases, relates to completion of a credential degree. The non re-enrolling student indicated the less student’s success condition because the non re-enrolling students is an early warning for student’s withdrawal, especially in ODL institution which is not applying drop-out regulation. Universitas Terbuka (UT) does not apply drop-out regulation for the students. Student could stop re-enrolling in one time and start to re-enroll again in another time or to decide to not re-enroll forever. Predicting the students’ re-enrollment pattern (may be) then was an attempt to diagnose a student condition for success and could be followed by specific and appropriate supports to prevent the student being withdrawal.

The purpose of this study was to predict student’s re-enrollment in UT.Since students of UT have greater flexibility to re-enroll or not in every semester, predicting student’s re-enrollment is not as simple as in conventional higher education environment. The student’s re-enrollment pattern was related to dropout or graduation rate. One of difficulties in predicting student’s re-enrollment in ODL was frequently associated with higher dropout rate. Simpson (2016) reported the dropout rates for a number of distance institutions compared with the UK full-time rate that distance institutions tend to have much higher dropout rates than the UK full-time average. In UT, the tendency of decreasing number of student’s re-enrollment in one semester to subsequent semesters was illustrated in Figure 1.

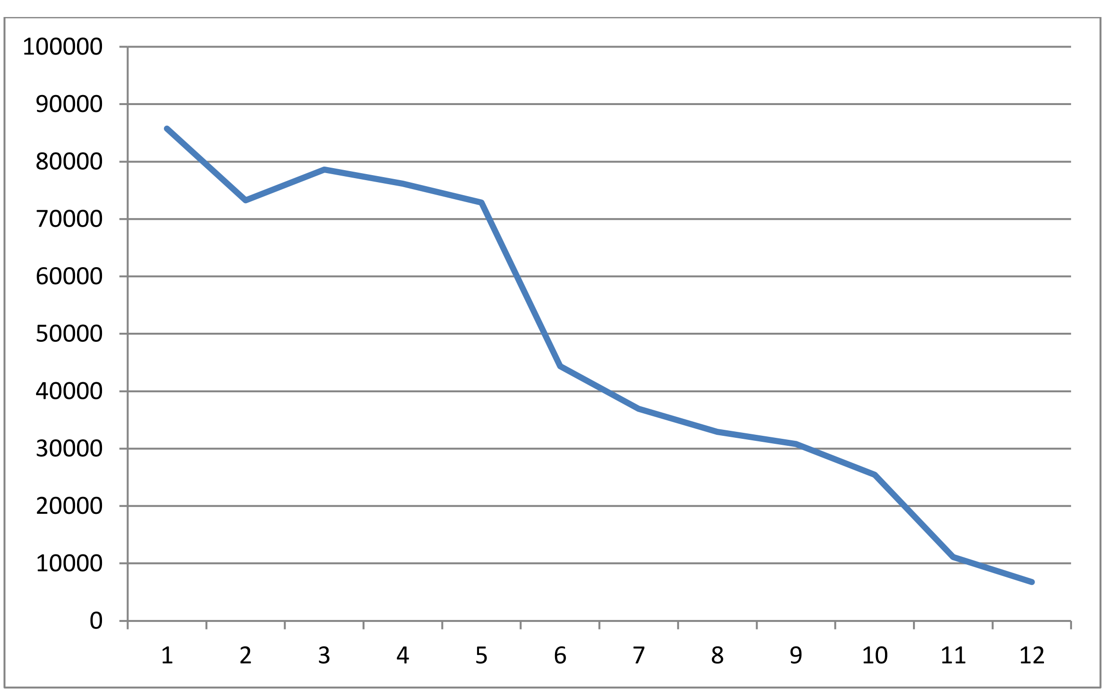

Figure 1. Decreasing Student’ Re-enrollment for 2009-2015

Literature Review

In this study, the term of re-enrollment was defined as condition that a student enrolling in one semester and enrolling again in the next semester sequentially. Non re-enrolling student phenomena in higher education could be accompanied to many terms in literatures. The term withdrawal, dropout, attrition often used to describe negatively the attribute of the non re-enrolling student, on the other hand, the terms of retention or persistence attributed for re-enrolling students (Draper, 2008). In most part of this study, re-enrollment is attributed to early sign of student success in open distance learning environment.This literature review would be focused on interaction in ODL as a theoretical framework to describe students success and some research results in predicting students success.

Dzakiria, Kasim, Mohamed, & Christopher (2013) asserted that the structure of ODL provides learners with the greatest flexibility. It provides control over time, place and pace of education. One important element of success factor for students attending the ODL program is the level of interactivity within the student-tutor-content dyads. Therefore, the writer believes that to improve ODL experience, decrease dropout rates and maintain success stories for ODL, tutors and all-important stakeholders in ODL must improve the provision of interaction and interactivity. Similarly, research regarding the importance of interaction in ODL conducted by Choi, Lee, Jung, & Latchem (2013) that indicated a lack of feedback from the instructors, heavy workload, and difficulties in studying at a distance were the main reasons for non re-enrollment. The learners' perceptions of the value of the degrees and their ages, gender, and educational backgrounds were also found to be significant factors in decisions not to re-enroll. The suggested solutions for reducing non re-enrollment include: a decrease in the number of required credit hours' study per semester; the provision of stronger social support; the introduction of a more flexible enrollment system; and better use of the available technology and infrastructure to help both students and instructors build stronger learning communities.

Describing why interaction in ODL is extremely important, Smart, Feldman & Ethington (2006), relying on Holland's theory, provide a theoretical linkage between variations in patterns of student success and students' learning experiences as well as their interactions with different academic environments. The first is that student success is a function of the fit or congruence between students' personality type and their chosen academic environments. The second is that student success is determined by the extent to which students learn the distinctive patterns of attitudes, interests, and abilities that are required, reinforced, and rewarded by their chosen academic environments, irrespective of the fit or congruence between students' personality types and their chosen academic environments. i.e., academic majors. In this report, the academic environment referred to academic majors, but in the light of the environment meaning it was easy to extend that the ODL environment should be included.

The above study results indicated the importance of interaction in a learning environment to prevent students from unsuccessful and therefore educators of the ODL strive to provide possibilities for interaction to motivate and enhance the chance of student success with their learning objectives in the ODL environment. Studies concerning of prediction of the student success were abounding in ODL higher education as well as in face-to-face universities.

Using an expanded person-environment fit (P-E fit) model for college students of science, technology, engineering, and mathematics (STEM), Le, Robbins &Westrick (2014) found that individual difference factors play the importance of integrating ability and interest constructs to fully understand college and career choice and persistence behaviors. One of the finding was that the relationship between ability and persistence is stronger for females than it is for males. Similarly, Godfrey&Matos-Elefonte (2010) studied the prediction of student success in college according to various student-level background and academic variables as well as school-level social and academic characteristics and demonstrated that characteristics at both levels play a role in the likelihood of reaching these goals. Also, Stephan, Davis, Lindsay, & Miller (2015) described the early college success of students, identifying measures in the state longitudinal data system that predict early college success, and examining the usefulness of those predictors.The study found that half the students achieved early college success by the composite of all three indicators, i.e.: enrolling in only nonremedial courses in the first semester, completing all attempted credits in the first semester, and persisting to a second year of college. The study also identified variables for student demographic characteristics, high school academic achievement, and behavior that might be related to (or predict) whether a student achieves success in the early college years. Again, analysis of predictive models of student success in college was reported by researchers at University of Maryland University College (2015) in identifying factors associated with student success based on students' GPAs and retention rates. The results showed that students' performance in their first semester at UMUC remains crucial in predicting re-enrollment, retention, and graduation. First term GPA may be an indicator of factors contributing to students' success, beyond academic abilities.

Predicting students success is particularly important for new students where the pre-course start information available is sometimes slight and withdrawal often occurs very early in a course. It suggests that, in such cases, statistical methods involving logistic regression analysis are more useful than questionnaires or tutors' opinions. Identifying students with low probability of success allows support to be targeted on them.

Resarch Method

The research was conducted in quantitative approach.The research problem in this study was to predict the re-enrollment of the students of UT in next semester based on student's participating in learning support services (face-to-face-tutorial, online tutorial), academic achievement (GPA, credits earned), and personal characteristics (age, gender, and marital status). The predicting variable was a dichotomous outcome: whether a student will re-enrollment or not in the next semester. The data analysis in this study was used binary logistic regression. The binary logistic regression was proposed as a statictical technique in the late 1960s and early 1970s,and it became routinely available in statistical packages in the early 1980s (Peng, Lee, & Ingersoll, 2002).

The logistic regression required only the dependent variable as a binary variable. The independent variables could be interval, ordinal, or categorical. However, to avoid too many empty cells and for convenient interpretations, in this study variables GPA, credits earned, and age were changed to categorical too, as shown in Table 1.

Table 1
The Independent Variables

Variables	Variable Description	Variable values
X_1	Marital status	• = 1 for married students • = 0 for unmarried students
X_2	Gender	• = 1 for male students • = 0 for female students
X_3	Age	• = 1 for student's age less than or equal to 25 • = 0 for students's age more than 25
X_4	Student's grade point average (GPA)	• = 1, if the GPA was more than 2.5 • = 0, if the GPA was les than or equal to 2.5
X_5	Student's credits earned	• = 1, if total of the credit earned was more than 12 • = 0, if total of the credit earned was less or equal to 12
X_6	Student's participating in face-to-face tutorial	• = 1, if the student participated • = 0, if the student did not participated
X_7	Student's participating in online tutorial	• = 1, if the student participated • = 0, if the student did not participated

The sampling technique used was simple random sampling (Borg & Gall, 1989). The sample size was 1195 new students of the first semester of 2015. The regression model was applied to predict the students' re-enrollment in the second semester of 2015. The model of the prediction used logistic regression that was represented by the following formula:

$$\log\left[\frac{p}{(1-p)}\right] = a + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5 + b_6X_6 + b_7X_7$$

where:

p – the probability of student's re-enrollment

a – a constant

b_i – coefficient of logistic regression for the variables i^{th} respectively

Results and Discussion

The 1195 students in the sample were categorized by predictors and the two groups of re-enroll and non re-enroll students as shown in Table 2. The proportion of unmarried and married students were varied in the two groups, the unmarried students proportion slightly increased in the non re-enrollment group. This situation was also appearing for gender variable where the female students increased from 60,95% in the re-enroll group to 77,29% in the non re-enroll group.

Table 2.
Description of Re-enroll and Non Re-enroll Groups

Predictors	**Re-enroll**		**Non Re-enroll**	
	Frequency	**Percent**	**Frequency**	**Percent**
Marital Status				
Unmaried	109	64.50	749	73,00
Maried	60	35.50	277	27,00
Gender				
Female	103	60,95	793	77,29
Male	66	39,05	233	22,71
Age				
Equal to or under 25	60	35,50	288	28,07
Over 25	109	64,50	738	71,93
Grade Point Average (GPA)				
Equal to or Less than 2.5	152	89,94	459	44,74
More than 2.5	17	10,06	567	55,26
Credits Earned				
Less than or equal to 12	108	63,91	191	18,62
More than 12	61	36,09	835	81,38
Taking part on Face-to-face Tutorial				
Yes	56	33,14	821	80,02
No	113	66,86	205	19,98
Taking part on Online Tutorial				
Yes	83	49,11	206	20,08
No	86	50,89	820	79,92

The percentage of students under or equal to 25 year old in the re-enroll a group (35.50%) somewhat higher than the non-re-ennroll group (28.07%). The re-enroll group dominated by students with a GPA less than 2.5 (89.94%) compared to the non-re-enroll group that was only 44.74%. Similarly, the re-enroll group of the student's credit earned tended to be dominated by more than 12 credits earned (81.38%) compared to the non-re-enroll group that was 36.09%. Taking part in face-to-face tutorial made the two groups were considerably different, 80.02% of the students in the re-enroll group took part in face-to-face tutorial compared to 33.14% in the non-re-enroll group. However, the situation was not applied to online tutorial where the student participating in online tutorial for the non-re-enroll group (49,11%) surpassed the re-enroll group (20,08%). The difference between the two groups in the Table 2 could be tested by chi square test to confirm whether the differences were significant or not and would be concluded which variables determined a student to re-enroll or not in the next semesters. However, the purpose of this study was predicting rather than comparing the two groups. Therefore, binary logistic regression was used to predict student's membership in the re-enroll group or the non-re-enroll group and the results were presented in Table 3.

Table 3
Variables in the Equation

Variables	B	S.E	Wald	Df	Sig.	Exp(B)
Marital Status	.433	.223	3.778	1	.052	1.543
Gender	.309	.221	1.962	1	.161	1.362
Age	-.032	.013	6.401	1	.011	.968
Grade Point Average (GPA)	1.038	.310	11.185	1	.001	2.823
Credits Earned	1.429	.213	45.130	1	.000	4.176
Taking part on F2F Tutorial	1.414	.270	27.489	1	.000	4.113
Taking part on Online Tutorial	-.148	.252	.343	1	.558	.863
Constant	1.250	.547	5.232	1	.022	3.491

Table 3 presents the statistical significance of the individual regression coefficients (βs) tested using the Wald Chi-square statistic. The test of the intercept was significant ($p<.05$) suggesting that the intercept should be included in the logistic regression equation model. The Hosmer & Lemeshow test of the goodness of fit suggested that the model was a good fit to the data as $p=0.841$ ($>.05$). The Nagelkerke's R^2 suggested that the model explains roughly 37% of the variation in the outcome. Finally, the model was classifying correctly 96.5% of students who enrolled and 40.8% of those who did not, for an overall success rate of 88.6%.

According to Table 3, age, grade point average, credits earned, and taking part on face-to-face tutorial were significant predictors for student's to re-enroll ($p<.05$). The last column, Exp (B) presents odds ratio for each of the predictors. The odds express the likelihood of an event occurring relative to the likelihood of an event not occurring (Park, 2013). In this case, the odds expressed the likelihood of the event of re-enrollment or non re-enrollment. The odds ratio (OR) is a comparative measure of two odds relative to different events (Park, 2013). The odds ratio is calculated by dividing the odds by other odds, for example in the Table 3, the odds ratio of the age variables was 0.968 (less than 1), it mean that the odds of the students' with age under or equal 25 was less than the students with age over 25 (see Table 1) to re-enrolled in the next semester by 0.968 times. The younger students tended slightly to less likely to re-enroll in the next semester than the older student. This finding was similar to the result of a research conducted by University of Maryland University College (2015) that age and marital status were associated with success at UMUC.

From the Table 3, a student with GPA more than or equal to 2.5 was 2.82 times more likely to re-enroll in the next semester than was a student with GPA below 2.5. GPA is associated with academic ability and previous academic success. It was not surprising that student's academic ability and previous success was associated with re-enrollment. The research finding suggested that academic ability was indeed related to persistence in college (Le, Robbins, & Westrick, 2014). This was comparable to Stephan, Davis, Lindsay, & Miller (2015) that having higher test scores and taking advanced coursework in previous school predicted indicators of early college success.

A student with credits earned more than or equal to 12 units is 4.17 times more likely to re-enroll than is a student with credits earned less than 12 units. Similar to GPA, the credits earned in one semester was a measurement of success and academic ability. Therefore, this finding was in line with the research results that academic ability was related to persistence in college (Le, Robbins, & Westrick; 2014; Stephan, Davis, Lindsay, & Miller; 2015).

A student took part in face-to-face tutorial was 4.11 times more likely to re-enroll than was a student who did not take part in face-to-face tutorial. Joining in a face-to-face tutorial was definitely making big difference for the student to re-enroll in the next semester. Face-to-face tutorial theoretically related to interaction in ODL. Choi, et al (2013) indicated that a lack of feedback from the instructors, heavy workload, and difficulties in studying at a distance were the main reasons for non re-enrollment. The face-to-face tutorial was considered as a means for the student to get a real feedback from the tutors and to make academic workload in self-directed learning of ODL envrionment to be less difficult.

However, the variable of student's taking part in face-to-face tutorial is supposed to correlate with the other predictor variables, especially with GPA and credits earned. Collinearity is a common problem for statistical modeling. In a linear regression model, dependencies among the covariates causes parameter estimates to be unstable (Haque, Jawad, Cnaan, and Shabbout; 2002). The calculation of Pearson correlation between student's taking part in face-to-face tutorial with GPA and Credits earned were 0.423 and 0.519 respectively and significant at the 0.01 level. But the relationship does not appear strong enough (Pearson's $r < 0.60$) to be considered a problem, usually values of $r = .8$ or more are cause for concern (Strand, Cardwallader, Firth, 2011).

Conclusion and Future Research

The core of ODL are access and success. The variable of age, GPA, credits earned, and participation in face-to-face tutorialwere significant predictors for student’s to re-enroll while the variable of marital status, gender, and participation in online tutorial were not. Predicting the students’ re-enrollment was an attempt to diagnose a student condition for success and could be followed by special and suitable supports to prevent the student of being withdrawal, such as:

- The younger students (below 25 years old) should receipt a special attention and urge them to interact with their peers, instructors, and institution. They need special program to support them, such as to get-together in social activities.
- The new student should be endorsed to take 12 or 15 credits in their first semester to get GPA more than or equal to 2.5.
- Participation in face-to-face tutorial is the best way for the students to get interaction with peers, instructor, and institution.

The future research of re-enrollment in ODL environment should be focused on effectivity of a special treatment to improve student persistence. The results of this research has given a light to implement the action. For instance, how to support the younger students, what academic work load will be suitable for some students with certain characteristics, and what kind of face-to-face tutorial will give the students the feeling of best interaction.

References

Borg, W. R., & Gall, M. D. (1989). *Educational research: An introduction (5th ed.)*. New York, NY: Longman. ISBN: 0-801-0334-6.

Choi, H., Lee, Y., Jung, I., & Latchem, C. (2013). The extent of and reasons for non re-enrollment: A case of Korea National Open University. *The International Review of Research in Open and Distance Learning, 14*(4). Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/1314.

Draper, S. (2008). Tinto's model of student retention. Retrieved from http://www.psy.gla.ac.uk/~steve/localed/tinto.htm

Dzakiria, H., Kasim, A., Mohamed. A.H., & Christopher, A.A. (2013). Effective learning interaction as a prerequisite to successful open distance learning (ODL): A Case study of Learners in the Northern State of Kedah and Perlis. *Malaysia Turkish Online Journal of Distance Education*, *14*(1). Retrieved from http://files.eric.ed.gov/fulltext/EJ1006252.pdf.

Godfrey, K. E. & Matos-Elefonte, H. (2010). *Key indicators of college success: Predicting college enrollment, persistence, and graduation*. New York, NY: College Board.

Haque A., Jawad A.F., Cnaan A., & Shabbout, M. (2002). *Detecting multicollinearity in logistic regression models: An extension of BKW diagnostic*. Proceeding of the 2002 Joint Statistical Meeting, American Statistical association, Statistical Computing Section.

Le, H., Robbins, S. B., & Westrick, P. (2014). Predicting student enrollment and persistence in college STEM fields using an expanded P-E fit framework: A large-scale multilevel study. *Journal of Applied Psychology*, *99*(5), 915-947.10.1037/a0035998

Park, Hyeoun-Ae. (2103). An introduction to logistic regression: From basic concepts to interpretation with particular attention to nursing domain. *Journal of Korean Academic Nursing*, *43*(2), 154-164. Retrieved from http://dx.doi.org/10.4040/jkan.2013.43.2.154.

Peng, C-Y., J., Lee, K.L., & Ingersoll, G.M. (2002). An introduction to logistic regression analysis and reporting. *The Journal of Educational Research*, *96*(1). Retrieved from http://www-psychology.concordia.ca/fac/kline/734/peng.pdf.

Sharma, J. (2015, September 18) . Open and distance learning – Access and success. *University World News*. Retrieved from http://www.universityworldnews.com.

Simpson, O. (2016). *Student Support Services for Success in Open and Distance Learning*. Available from http://www.cemca.org.in.

Smart, J.C., Feldman, K.A., & Ethington, C.A. (2006). *Holland's theory and patterns of college student success*. Available from http://nces.ed.gov/npec/pdf/smart_team_report.pdf.

Stephan, J. L., Davis, E., Lindsay, J., & Miller, S. (2015). *Who will succeed and who will struggle? Predicting early college success with Indiana's Student Information System*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. Retrieved from http://ies.ed.gov/ncee/edlabs.

Strand, S., Cardwallader, S., Firth, D. (2011). *Using statistical regression methods in education research.* Available from http://www.restore.ac.uk/srme/www/fac/soc/wie/research-new/srme/modules/index.html.

University of Maryland University College. (2015). *Predictive analytics for student success: Developing data-driven predictive models of student success*. Retrieved from https://www.umuc.edu/visitors/about/ipra/upload/developing-data-driven-predictive-models-of-student-success-final.pdf.