



Learning Interest Modelling of Poliwangi Students to Learn Mathematics Engineering Through MOOCs Using Dummy Regression

Ika Yuniwati¹, Aprilia Divi Yustita², Siska Aprilia Hardiyanti³, I Wayan Suardinata⁴

^{1,2,3,4}Politeknik Negeri Banyuwangi
Email: ika@poliwangi.ac.id

ABSTRACT

MOOCs is a learning system in the form of online courses that is massive and open to allow participants to enjoy unlimited content and can be accessed via the web. Mathematical techniques taught using MOOCs which will be developed in the following year are expected to be liked by students. The purpose of this study was to determine student interest in studying MOOCs. This study uses a dummy regression model on learning hours in each category. Dummy regression is considered a suitable model because dummy regression can quantify qualitative data. Qualitative data were obtained from a questionnaire distributed to 240 students. The questionnaire contains indicators of student MOOCs interest, including cognitive, affective, and psychomotor interests. The result of this study is the amount of time studying mathematics influenced by students' interest in learning mathematics through MOOCs by 60.7%, and the rest 39.3% is influenced by other factors. The model is $Y_i = 1,562 + 4,729 D_1 + 1,461 D_2 + \varepsilon_i$. So it can be concluded that the interest of students who want to study mathematics through MOOCs is the highest with an average student learning hours of 4,729 minus 1,562 equal to 3,647 hours.

Keywords: Dummy Regression Model; Learning Interest; MOOCs

INTRODUCTION

Massive Open Online Courses (MOOCs) can be qualified as revolution education has begun to grow and become popular today. Individuals can get training in the areas needed and developed with educational training that is open to all students throughout the world [1]. MOOCs caters to a large number of students and provides a combination of open online courses, short video lectures, automated conversations, quizzes, peer and self-talk, and student collaboration through discussion forums. There are various kinds of MOOCs designed according to the level of thinking since 2016 [2]. Massive open online courses (MOOCs) contribute significantly to individual empowerment because they can help people learn about various topics [3].

The goal of MOOCs is the best learning resource and new ways of learning in the classroom. Learning can help students learn fully, work together, and is also supported by expert guidance [4]. The use of MOOCs in learning already exists in various countries. Singapore MOOCs can reduce university and university level tuition fees improve community access to such courses. They also provide skills and job training for community members [5]. This contradicts the results of research which state that MOOCs can provide new forms of learning through technology and save significant costs for education [6]. In Portugal, there is a study that states there is a relationship between interest in educational success [7]. Other studies add learning competencies that can influence participation, perseverance, and sustainability [8]. A good formal education system supports the students involved. factors that affect this performance according to

a survey conducted at the Nigerian Private University (Redeemer University) namely hours of study [9]. While the learning domain that must be learned in learning can be categorized as the cognitive domain (knowledge), the psychomotor domain (skills), and the affective domain (attitude) according to Bloom [10]. Because this research analyzed interests published in the domain which is an indicator that shows interest in MOOCs. The analysis used dummy analysis. Dummy analysis can be used to predict interest. The research that has been done is the students' interest in soap [11]. The Dummy regression model is also used for the performance of students majoring in mathematics FMIPA [12]. Dummy variables have often been used in strategy research to study the effects of categorical variables [13,14]. The advantage of using these puppet variables, variables 1 and 0 is that they can be questioned and interpreted as the resulting regression estimates [15].

METHODS

Multiple Linear regression analysis

The Dummy regression analysis is a double linear regression analysis whose variable is qualitative. So before the use of dummy regression analysis, it must first be understood the analysis of a double linear regression [12] following
On each observation, represented the i^{th} bservation, applies the equation

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_p X_{pi} + \varepsilon_i \quad (1)$$

System equations (1) can be written in the form of a matrix, by defining each matrix into the following matrix:

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}; X = \begin{bmatrix} 1 & X_{11} & X_{12} & \dots & X_{1k} \\ 1 & X_{21} & X_{22} & \dots & X_{2k} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & X_{n1} & X_{n2} & \dots & X_{nk} \end{bmatrix}; \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}; \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad (2)$$

or equations (2) can be written in the form of another matrix as follows :

$$Y = X\beta + \varepsilon \quad (3)$$

Based on the assumptions above $\varepsilon_i \sim N(0, \sigma^2)$, Then the equation (1) can be written in the form of expectation value:

$$E(Y_i) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ik} \quad (4)$$

Estimation parameters

The estimation of the parameters can be obtained using the smallest quadratic method so that the equation (4) can be written in a matrix form :

$$\hat{\beta} = (X'X)^{-1}X'Y \quad (5)$$

Hypothesis testing was conducted to test the overall regression Parameter. Overall test of the regression parameters as follows:

$$H_0: \beta_0 = \beta_1 \dots = \beta_k = 0$$

$$H_1: \text{at least one } \beta_j \neq 0$$

The Sum of Squares Total (SST) is The Total of Sum Squared Regression (SSR) and The Sum of Squared Error (SSE), or it can be written:

$$SST=SSR+SSE \tag{6}$$

Test statistics which used are F test statistics:

$$F = \frac{SSR/DFR}{SSE/DFE} = \frac{MSR}{MSE}$$

Explanation

DFR = Degree of freedom in regression

DFE = Degree of freedom in error

MSR = Mean Sum of Squares Regression

MSE = Mean Sum of Squares Error

H_0 rejected if $F_0 > F_{(\alpha,k,n-k-1)}$

By minimizing the number of squared errors, it is obtained:

$$SSE = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \tag{7}$$

$$SST = \sum_{i=1}^n (Y_i - \bar{Y}_i)^2 \tag{8}$$

From equation (6), (7), and (8) obtained

$$SSR = \hat{\beta}'X'Y' - \frac{(\sum_{i=1}^n Y_i)^2}{n} \tag{9}$$

When Variable a free variable is inserted one by one gradually into a regression equation, it is performed a sequential F test [16]

Hypothesis testing for partial regression coefficient parameters

The F test is used to determine the effect of the independent variable on the dependent variable simultaneously. After the F test is carried out, the T-test or partial regression test is carried out. This test is used to determine the effect of each independent variable on the dependent variable. Partial regression hypothesized will be testing

$H_0 : \beta_j = 0$

$H_1 : \beta_j \neq 0$

Test statistics: $t_{hit} = \frac{\hat{\beta}_j}{Se(\beta_j)}$

(10)

H_0 Rejected if $|t_{hit}| > t_{(\frac{\alpha}{2},n-k-1)}$

Coefficient of determination

After knowing the effect simultaneously and the effect of each independent variable. The next step of analysis is to find the percentage of the independent variables as a whole to the dependent variable. Multiple coefficients of determination R^2 measures the proportion of total diversity in the Y-free variables that can be explained by the regression equation model together. Size of regression coefficient determined by the formula:

$$R^2 = \frac{SSR}{SST} \tag{11}$$

Then from the independent variable and the dependent variable, the model is determined. There are many ways to build a regression model whose free variables contain. Variable A qualitative variable, one of which is using a variable doll or commonly called a dummy variable. Dummy variables are used as an attempt to see how

the classifications in the sample affect the estimation parameters. Variables dummy also tries to make the quantification of Qualitative variables. For example, if you want to estimate Variable. The value of a variable Y is influenced by one Variable quantitative variable (X) and one Variable qualitative free variable that has two categories, such as Category 1 and Category 2. The dummy model of the example is

- a. $Y = a + bX + cD1$ (Dummy Intercept Model)
- b. $Y = a + bX + c(D1X)$ (Dummy Slope Model)
- c. $Y = a + bX + c(D1X) + dD1$ (Dummy Intercept and Slope Model)

This research used Dummy Intercept Model or Sympel Dummy Regression Model by Modupe by the formula [9]

$$Y_i = b_0 + b_1 Z_i + e_i$$

b_0 = intercept

b_1 = Regression coefficient

$z_i = 1$ If the unit to I is a group that is

= 0 If the unit to I as the reference group

RESULTS AND DISCUSSION

Data Presentation

Respondents, in this case, were 240 students. These students have been given knowledge about the MOOCs that was developed. Students are given a questionnaire totaling 20 questions. The questionnaire was designed to contain indicators of interest in the cognitive, affective, and psychomotor domains. After completing the questionnaire then make groups of them. It is namely students into students who are interested in learning mathematics through MOOCs (A), students who do not like to learn mathematics through MOOCs (B), and students who do not want to learn mathematics through MOOCs (C). The grouping scores can be seen in Table 1.

Table 1. Student Interest Grouping Score

Category	Score Interval
A	60-80
B	40-59
C	20-39

The data in the research has been categorized as in Table 1 then it was changed to dummy variables. The reference category is chosen. It is C Category. So that Category C becomes D0, Category A becomes D1, and Category C becomes D2.

Test of Significance and Coefficient of Determination

The Significance Test is divided into 2 tests, namely the F-Test to find out the significance simultaneously and the T-Test to find out the partial significance. F-Test is used to test a regression that is by testing hypotheses that involve more than one coefficient. The F-Test can also be used to test the linearity of a regression equation. It can also be used to see the effect between independent and dependent variables. F-Test on the results of this study can be seen in Table 2.

Tabel 2. Result of F test

<i>F-statistic</i>	F	<i>p-value</i>
183,313	2	0.000

Table 2 indicate that the value of the F-Test is 183.313 with a p-value of 0,000, it can be stated that the model or independent variable factors, in this case, the variable of interest in studying Engineering Mathematics through MOOCs affects the dependent variable, the amount of time studied

The T-Test can be used to test hypotheses about individual coefficients, the T-Test is also often called a partial test. The results of the T-Test can be seen in Table 3.

Tabel 3. Result of T test

<i>Coefficient</i>	<i>t-value</i>	<i>sig</i>
D0	6.282	0.000
D1	16.040	0.000
D2	5.257	0.000

Based on Table 3 above, it is known that all predictor variables significantly influence the model. It can be seen sig value which has a value less than 0,05. So, the regression coefficient in each category produced on the variable significantly influences the number of learning hours.

The coefficient of determination is used to find out how much influence is dependent on the independent variable. It can be seen in Table 4.

Tabel 4. Table of effect from the amount of time studying in mathematics towards students interest to MOOCs

Model	R	R ²
1	0.779 ^a	0.607

a. Predictors: (Constant), D2, D1

b. Dependent Variable: the amount of time studying mathematics

Table 4 reflected the correlation between the amount of time studying mathematics and students's interest in learning mathematics through MOOCs. R-value is 0,779. This value shows that their correlation is strong. Besides, R² shows a value of 60.7%, This gives the meaning of the amount of time studying mathematics influenced by students' interest in learning mathematics through MOOCs by 60.7%, and the rest 39.3% is influenced by other factors.

Interpretation of Dummy Regression Model

The data in this model uses the results of the data which are categorized into 3 dependent variables. It can be seen in Table 5.

Tabel 5. The estimated Learning Interest through The Amount of Time Studying Mathematics in 3 Categories

Model	Unstandardized Coefficient		Sig.
	B	Std. Error	
1 (Constat)	1,562	0,249	0,000
D1	4,729	0,295	0,000
D2	1,461	0,278	0,000

Table 5 shows that they are students who do not want to learn mathematics through MOOCs (D0), students who are interested in learning through MOOCs (D1), students who do not like to learn through MOOCs (D2) and 1 independent variable the average hours of study (Y) are dummy and a linear regression model is obtained

$$Y_i = 1,562 + 4,729 D1 + 1,461 D2 + \varepsilon_i$$

The interpretation of the regression model above is the average study time for students who do not want to learn engineering mathematics by 1,562 hours. The average difference in hours of study for students interested in learning engineering mathematics through MOOCs with students who do not want to learn engineering mathematics through MOOCs is 4,729 hours or in other words the interest of students who do not want to study engineering mathematics through MOOCs is lower than the interest of students interested in learning engineering mathematics through MOOCs. The average study time for students who interested in learning engineering mathematics by 3,467 hours. The average difference in hours of study for students who do not like to study engineering mathematics through MOOCs with students who do not want to study engineering mathematics through MOOCs is 1,461 hours or in other words the interest of students who do not likes to study engineering mathematics through MOOCs is lower than that of students who do not want to study engineering mathematics through MOOCs. The average study time for students who do not like study engineering mathematics by 0,101 hours or equals to 6 minutes.

CONCLUSIONS

The results of this study are interest of students who do not want to study engineering mathematics through MOOCs is lower than the interest of students who are interested in learning engineering mathematics through MOOCs. Moreover, the interest of students who do not want to study engineering mathematics through MOOCs is lower than the interest of students who do not like to study engineering mathematics through MOOCs.

REFERENCES

- [1] H. Bicen, "Determining the effect of using social media as a MOOCs tool," *Procedia Comput. Sci.*, vol. 120, pp. 172–176, 2017, doi: 10.1016/j.procs.2017.11.225.
- [2] C. Wrigley, G. Mosely, and M. Tomitsch, "Design Thinking Education: A Comparison of Massive Open Online Courses," *She Ji*, vol. 4, no. 3, pp. 275–292, 2018, doi: 10.1016/j.sheji.2018.06.002.
- [3] M. Aparicio, T. Oliveira, F. Bacao, and M. Painho, "Gamification: A key determinant of massive open online course (MOOCs) success," *Inf. Manag.*, vol. 56, no. 1, pp. 39–54, 2019, doi: 10.1016/j.im.2018.06.003.
- [4] J. Ma, J. Zheng, and G. Zhao, "The Applicable Strategy for the Courses Alliance in Regional Universities Based on MOOCs Platform," *Procedia - Soc. Behav. Sci.*, vol. 176, pp. 162–166, 2015, doi: 10.1016/j.sbspro.2015.01.457.
- [5] V. Lim, L. Wee, S. Ng, and J. Teo, "Massive Open and Online Courses (MOOCs) and Open Education Resources (OER) in Singapore," *J. Southeast Asian Educ.*, vol. 1, pp. 1–13, 2017.
- [6] R. Pollack Ichou, "Can MOOCs reduce global inequality in education?," *Australas. Mark. J.*, vol. 26, no. 2, pp. 116–120, 2018, doi: 10.1016/j.ausmj.2018.05.007.
- [7] P. Goulart and A. S. Bedi, "The Impact of Interest in School on The Impact of Interest in School on Educational Success in Portugal," Bonn, Germany, 2011.
- [8] W. Abeer and B. Miri, "Students' Preferences and Views about Learning in a

- MOOCS," *Procedia - Soc. Behav. Sci.*, vol. 152, pp. 318–323, 2014, doi: 10.1016/j.sbspro.2014.09.203.
- [9] O. D. Modupe, "A Dummy Variable Regression On Students ' Academic Performance," *Transnatl. J. Sci. Technol.*, vol. 2, no. 6, pp. 47–54, 2012.
- [10] M. E. Hoque, "Three domains of learning - cognitive, affective, and psychomotor," *J. EFL Educ. Res.*, vol. 2, no. 2, pp. 45–52, 2016, [Online]. Available: <https://www.mendeley.com/catalogue/three-domains-learning-cognitive-affective-psychomotor-second-principle/>.
- [11] D. Sipahutar, P. Bangun, and U. Sinulingga, "Analisa Faktor Ketertarikan Mahasiswa Terhadap Produk Sabun Mandi," *Saintia Mat.*, vol. 1, no. 2, pp. 175–185, 2013.
- [12] N. Amalita and Y. Kurniawati, "Model Regresi Dummy dalam Memprediksi Performansi Akademik Mahasiswa Jurusan Matematika FMIPA UNP," in *Prosiding Semirata FMIPA Universitas Lampung*, 2013, pp. 387–391.
- [13] P. S. L. Yip and E. W. K. Tsang, "Interpreting Dummy Variables and Their Interaction Effects in Strategy Research," *Strateg. Organ. J.*, vol. 5, no. 1, pp. 13–30, 2007, doi: 10.1177/1476127006073512.
- [14] M. Venkataramana, M. Subbarayudu, M. Rajanis and K. N. Sreenivasulu, "Regression Analysis with Categorical Variables," *International Journal of Statistics and Systems*, vol. 11, no. 2, pp. 135–143, 2016
- [15] I. C. A. Oyeka and C. H. Nwankwo, "Use of Ordinal Dummy Variables in Regression Models," *IOSJRM*, vol. 2, no. 5, pp. 1–7, 2012.
- [16] N. . Draper and S. H., *Applied Regression Analysis*, no. 48. 1981.