

## Application of Binary Logistic Regression Analysis Method with Spss Statistics 22.0 in Predicting Factors Affecting The Long of Study

Dian Yulis Wulandari, Ismaraidha, Ocdy Amelia  
Universitas Pembangunan Panca Budi <sup>1-3</sup>, Indonesia<sup>1-3</sup>

[dianyulis@dosen.pancabudi.ac.id](mailto:dianyulis@dosen.pancabudi.ac.id)

### Abstract.

### Keywords:

### Keywords Abstract

Length of the  
study period,  
Binary  
Logistics  
Regression,  
Intelligence

Binary logistic regression is a statistical method that analyzes the relationship between response or independent variables that have two or more categories with explanatory or dependent variables on a dichotomous scale (2 categories). Determination of the binary logistic regression model in addition to the usual calculation method can also use the help of IBM SPSS 22.0. The purpose of this study is to provide information on the stages in managing data related to the Binary Logistics Regression Model using IBM SPSS 22.0. The factors that affected how long the study period was, were the examples of cases that were examined in this study. Academic aptitude, motivation, environment, lecturers, physical and mental health, interests and talents, learning styles, intelligence (IQ), and study time are the response variables (independent) that have an impact on the length of the study period in this study. The sample for this study, which employed the sample clustering method, consisted of 54 respondents from the Panca Budi Development University's faculty, staff, and students. A questionnaire-based survey is the technique used. According to the findings of this study's analysis, the variable belonging to the category of intelligence (IQ) has a tenuous relationship with the dependent variable over the course of the investigation. The dependent variable is simultaneously impacted by the independent variable by 41.5%, with other factors influencing the remaining percentage. The relationship between the independent and dependent variables could be predicted with 79.6% accuracy using calcification.

### Introduction

Education is an important thing in human life that can be obtained for everyone, starting from the young to the old. Education is a learning process regarding knowledge and skills aimed at educating and developing one's potential. This proves that all humans have the right to education and to develop both intelligence and skills.

The young generation is the front line in building the nation and is the next generation of the nation. Preparing a competent young generation is the same as preparing the frontline who are also human resources in a country. A great country will make education its main priority because education is the gateway for humans to lead to a developing and prosperous future.

One of the indicators in making reliable human resources is education, especially higher education, namely lectures. The quality of education, the quality of teachers (lecturers), student

intelligence, and higher education facilities are among the supporting factors. Not only that, the competence of students in their roles and participation to commit to completing the study period on time in higher education can also be used as a factor in producing reliable human resources.

Graduation is a moment of confirmation and inauguration ceremony for a student who has completed his studies at a university. Higher education is a formal education that has the duty and responsibility in preparing students who have graduate competencies by the goals of national education and the views and ideals of the founding fathers of the nation. Where the college will prepare candidates for bachelor, master, and doctorate who are reliable and have the best expertise in their respective fields.

In achieving a bachelor's degree, Master's and Doctoral degrees require a predetermined normal study time. The study period for the Bachelor (S1) program is for 4 (Four) Years, the Study Period for the Post-Master's Program (S2) is for 2 (Two) Years, and the Doctoral Program (S3) is for 4 Years. However, in practice so far, there are still many students who cannot complete their studies on time. The factors that cause the inaccuracy of student graduation time can be sourced from external factors or internal factors.

The length of the study period in the undergraduate program (S1) at Universitas Pembangunan Panca Budi is divided into two, namely less than 4 (four) years and more than or equal to 4 (four) years and the quality of graduates can be measured based on the Grade Point Average (GPA). measure based on the Grade Point Average (GPA). The success of students in taking education is influenced by several factors that influence each other. The study period is one of the factors measuring student success in the field of education. Students who complete their study period less than or equal to the normal time that has been determined is one of the assessments of students who are said to be successful in the field of education. On the other hand, students who complete their study period longer than the normal time that has been determined are considered less successful. Based on the things described above, the researchers are interested in analyzing the Application of the Binary Logistics Regression Model with SPSS Statistics 22.0 in Predicting Factors Affecting the Length of the Study Period.

The length of the student's study period in carrying out education in higher education is influenced by internal factors and external factors, namely:

#### **A. Internal Factors**

##### **1. Physiological Factors (Physical and Mental Conditions)**

According to Ginting in his book *Tips for Studying in Higher Education in 2005* that maintaining physical and mental conditions is one of the efforts to maintain learning abilities.

##### **2. Psychological Factors**

###### **a. Motivation**

Motivation is a psychological condition that encourages someone to do something. A motivated person is a person who tends to him to strive to achieve goals to satisfy all his needs. According to Ginting (100:2005), a person's motivation to learn will determine the amount of learning effort made.

b. Interest

Interest means the tendency of the heart (desire, liking) towards something (Ginting, 2005).

c. Talent and Intelligence

According to Hilgard quoted by Slameto (2010: 57), talent or aptitude is "the capacity to learn". This means that talent is the ability/competence to learn.

d. Study Methods and Habits

Learning aims to gain knowledge, and improve the quality and quantity of human behavior, such as improving skills, understanding, thinking power, skills, attitudes, knowledge, habits, and other abilities improvement (Thursan Hakim (2005: 52)).

## **B. External Factors**

### 1. Study Facilities

Facilities are everything in the form of objects or money that facilitates and facilitates the implementation of a particular business. Supporting facilities for the Education Area, both inside and outside the study room, need to be provided to support student activities in the lecture process and facilitate and expedite an activity both inside the study room and outside the study room or university area.

### 2. Teaching and Learning Process

In lectures at universities, there is an interaction between lecturers and students, where the lecturer will present teaching materials containing the sciences, and students will capture with enthusiasm every material or knowledge being taught. This interaction is referred to as the teaching and learning process. It takes methods and expertise by lecturers to be able to make this teaching and learning process interesting and accepted by students.

### 3. Study Time

According to Ginting (2005: 84), each student has a different study time, depending on what goals he will achieve. Students who have high ideals and life goals usually have a long study time because the student understands. Those high ideals will require a great struggle and are not easy.

### 4. Environment (Family, Community, and Campus)

According to Ginting (2005: 10) campus is an educational environment that is also a social environment. A healthy campus environment is not created just like that, but there is a need for guidance and cooperation between the entire academic community.

## **Regression Analysis**

Regression analysis is an analysis that aims to determine the effect of one variable on other variables. The simplest regression model is a simple linear regression model in the form of an equation (Suyono, 2015):

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

Information :

$Y$  = dependent variable (predicted value)bound (predicted value)

$X$  = independent variable

$\beta_0$  = Constant

$\beta_1$  = Regression coefficient (increase or decrease the value

$\varepsilon$  = Random error

## Binary Logistic Regression Analysis

Binary logistic regression is a statistical method that can analyze the relationship between several independent variables/predictors (X) on an ordinal or nominal scale with one or more dependent/response (Y) variables that are dichotomous (variables that divide respondents into 2 categories) (Hosmer and Lemeshow, 1989). For example: Passed (1) and Failed (0), Satisfied (1) and Dissatisfied (0), Buying (1) and Not Buying (0). The advantage of binary logistic regression does not assume the normality of the model error. Furthermore, the predictor variables in the binary logistic regression model can be mixed in scale, namely: they can be continuous and discrete variables. The stages of analysis in binary logistic regression are as follows:

1. The Binary Logistics regression equation model is as follows:

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_1 x_2 + \dots + \beta_k x_k)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_1 x_2 + \dots + \beta_k x_k)}$$

$\pi(x)$  (odds) adalah is the probability of a “successful” event, i.e.  $Y = 1$  with a probability value, of  $0 \leq \pi(x) \leq 1$ .

2. To obtain a linear form in logistic regression, it is necessary to determine the Logit Transform as follows:

$$g(x) = \ln \left[ \frac{\pi(x)}{1 - \pi(x)} \right] = \beta_0 + \beta_1 x_1 + \beta_1 x_2 + \dots + \beta_k x_k$$

The Logit function  $g(x)$  is a linear function in its parameters.

$\beta_0 + \beta_1 x_1 + \beta_1 x_2 + \dots + \beta_k x_k = \text{MLE (Maximum Likelihood Estimation)}$ .

3. Parameter Test

- a. Simultaneous Test

$H_0: \beta_1 = \beta_2 = \dots = \beta_k$  (There is no effect of the explanatory variable on the response variable)

$H_1: \text{At least there is one } \beta_j \neq 0$  (At least there is one explanatory variable that affects the respondent variable);  $j = 1, 2, 3, \dots, k$

- b. Partial Test

$H_0: \beta_j = 0$  (There is no influence between the  $j$ th explanatory variables on the response variable)

$H_1: \beta_j \neq 0$  (There is an effect of the  $j$ th explanatory variable on the response variable)

## Wald's Test Statistics:

$$W = \left( \frac{\beta_j}{SE(\beta_j)} \right)^2 \quad \rightarrow H_0 \text{ rejected if } W > X_{\alpha;1}^2 \text{ or p-value} < \alpha$$

4. Binary Logistics Regression Model Interpretation (Odds Ratio)

$$\Psi = \frac{\left[ \frac{\pi(1)}{\{1-\pi(1)\}} \right]}{\left[ \frac{\pi(0)}{\{1-\pi(0)\}} \right]} = \exp \exp (\beta_j)$$

Information :

- Odds Ratio ( $\Psi$ ) = 1  
Individuals with a value of  $x = 1$  have the same tendency as individuals with a value of  $x = 0$  to experience  $Y = 1$
- Odds Ratio ( $\Psi$ ) > 1  
Individuals with a value of  $x = 1$  have a greater tendency than individuals with a value of  $x = 0$  to experience  $Y = 1$
- Ratio Odds ( $\Psi$ ) < 1  
Individuals with a value of  $x = 1$  have a smaller tendency than individuals with a value of  $x = 0$  to experience  $Y = 1$

### Application Of Binary Logistic Regression Analysis Method With Spss Statistics 22.0

In this study used a case example with 54 respondents who are Lecturers, Students, and Employees at the Panca Budi Development University who have completed their undergraduate studies (S1).

The method used in this research was a survey method. Primary data collection was carried out by distributing questionnaires to respondents in October 2022. Data processing in this study used SPSS 22. The sampling technique used in this study was cluster sampling, namely, the sampling technique was carried out based on certain groups. The research variables and the number of questionnaires distributed are as follows:

Table 1. Data on the Distribution of Questionnaires

Variable	Category	Amount	Percentage (%)
<b>Total Sample</b>		<b>54</b>	
Study Period	> 4 Years	21	38,9
	≤ 4 Years	33	61,6
Academic Ability	High	30	55,6
	Medium	22	40,7
	Low	1	3,7
Motivation	High	33	61,1
	Medium	20	37
	Low	1	1,9
Environmental Factors (Family, Community, College)	High	36	66,7
	Medium	16	29,6
	Low	2	3,7
Lecturers who have extensive knowledge, simplify the process	High	31	57,4
	Medium	22	40,7
	Low	1	1,9
Physical Condition	Unhealthy	0	0

	Healthy	54	100
Mentally	Not Good	2	3,7
	Good	52	96,3
Student Interest	High	26	48,1
	Medium	25	46,3
	Low	3	5,6
Student Talent	High	18	33,3
	Medium	33	61,1
	Low	3	5,6
Learning Method	High	27	50
	Medium	25	46,3
	Low	1	3,7
Intelligence (IQ)	High	29	53,7
	Medium	23	42,6
	Low	2	3,7
Study Time	< 7 Hours	19	35,2
	≥ 7 Hours	35	64,8

Table 1 above, shows that 61.6% of the 54 respondents of Lecturers, Employees, and Students at the Panca Budi Development University have completed their undergraduate study period (S1) less than 4 (Four) years. 55.6% of respondents have the high academic ability. 61.1% of respondents have high motivation. 66.7% of respondents chose Environmental Factors (Family, Community, and Campus) to influence the study period. 57.4% of respondents choose Lecturers who have broad knowledge, simplify the process, and have high influence. 100% of respondents are in good health in carrying out undergraduate studies (S1). 96.3% of respondents have a good mentality. 48.1% of respondents have a moderate interest. 61.1% of respondents have a moderate level of talent. 50% of respondents chose the learning method given by a high-influenced lecturer. 53.7% of respondents with high intelligence and 64.8% of respondents have study time greater than or equal to 7 hours in 1 (one) week.

The steps in analyzing Binary Logistics Regression are as follows:

1. Display of IBM SPSS 22.0 there is 2 Sheets, namely: Data View and Variable View. First of all, make a list of independent variables and dependent variables in the variable view.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	X1	Numeric	8	2	Kemampuan A...	None	None	8	Right	Scale	Input
2	X2	Numeric	8	2	Motivasi	None	None	8	Right	Scale	Input
3	X3	Numeric	8	2	Faktor Lingkun...	None	None	8	Right	Scale	Input
4	X4	Numeric	8	2	Dosen	None	None	8	Right	Scale	Input
5	X5	Numeric	8	2	Kondisi Fisik	{,00, Sakit}...	None	8	Right	Nominal	Input
6	X6	Numeric	8	2	Mental	{,00, Tidak ...	None	8	Right	Nominal	Input
7	X7	Numeric	8	2	Minat	None	None	8	Right	Scale	Input
8	X8	Numeric	8	2	Bakat	None	None	8	Right	Scale	Input
9	X9	Numeric	8	2	Metode Pembel...	None	None	8	Right	Scale	Input
10	X10	Numeric	8	2	Kecerdasan (IQ)	None	None	8	Right	Scale	Input
11	X11	Numeric	8	2	Waktu Belajar	{,00, < 7 Ja...	None	8	Right	Nominal	Input
12	Y	Numeric	8	2	Lama Masa Studi	{,00, > 4 Ta...	None	8	Right	Nominal	Input

Figure 1. Display of Independent and Bound Variable Categories in "Variable View"

2. Input the value of the questionnaire results in the "Data View"

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	Y
1	3.00	2.00	1.00	2.00	2.00	2.00	3.00	2.00	2.00	3.00	2.00	1.00
2	2.00	2.00	3.00	3.00	2.00	1.00	3.00	2.00	2.00	2.00	2.00	2.00
3	3.00	3.00	3.00	3.00	2.00	2.00	3.00	3.00	3.00	3.00	2.00	2.00
4	2.00	2.00	2.00	1.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	1.00
5	2.00	2.00	3.00	3.00	2.00	2.00	2.00	3.00	2.00	3.00	2.00	1.00
6	3.00	2.00	3.00	3.00	2.00	2.00	2.00	1.00	2.00	3.00	1.00	2.00
7	3.00	2.00	2.00	3.00	2.00	2.00	2.00	2.00	3.00	2.00	2.00	2.00
8	3.00	3.00	3.00	2.00	2.00	2.00	2.00	3.00	3.00	3.00	2.00	2.00
9	2.00	2.00	3.00	3.00	2.00	2.00	2.00	2.00	2.00	3.00	2.00	2.00
10	3.00	3.00	3.00	3.00	2.00	2.00	3.00	2.00	3.00	3.00	2.00	1.00
11	3.00	3.00	2.00	3.00	2.00	2.00	2.00	3.00	3.00	3.00	2.00	2.00
12	3.00	3.00	3.00	2.00	2.00	2.00	3.00	3.00	3.00	3.00	2.00	1.00
13	2.00	2.00	3.00	2.00	2.00	2.00	2.00	2.00	3.00	3.00	2.00	1.00
14	1.00	3.00	2.00	2.00	2.00	2.00	3.00	2.00	2.00	3.00	2.00	2.00
15	2.00	3.00	1.00	2.00	2.00	2.00	3.00	3.00	2.00	3.00	1.00	1.00
16	3.00	3.00	3.00	3.00	2.00	2.00	2.00	2.00	2.00	2.00	1.00	1.00
17	3.00	2.00	3.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	1.00	1.00
18	2.00	3.00	3.00	3.00	2.00	2.00	2.00	2.00	3.00	2.00	1.00	2.00
19	2.00	2.00	2.00	2.00	2.00	2.00	2.00	3.00	3.00	3.00	2.00	1.00
20	3.00	3.00	3.00	3.00	2.00	2.00	2.00	3.00	3.00	3.00	1.00	2.00
21	2.00	3.00	3.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00

Figure 2. Display of questionnaire results data on "Data View"

3. Next select Analyze, select Regression, and select Binary Logistics.

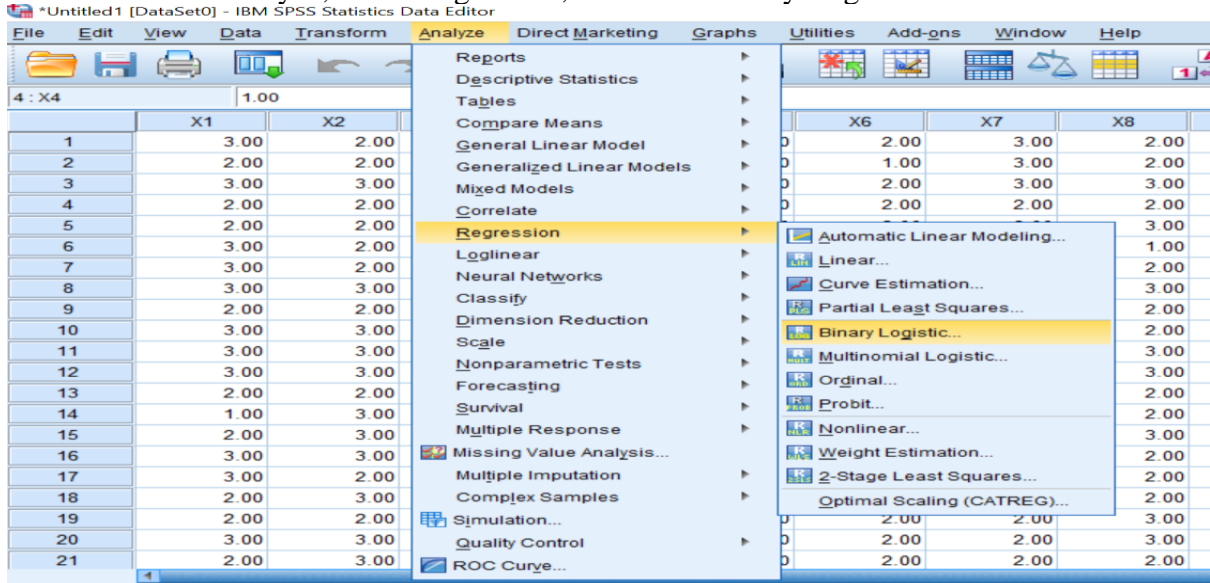


Figure 3. Stages in Logistics Regression Analysis

4. Move the dependent variable (Y) into the dependent column and the independent variable (x) into the covariate's column. Then select Options and put a checkmark on "Statistics and Plots", then select 'OK'.

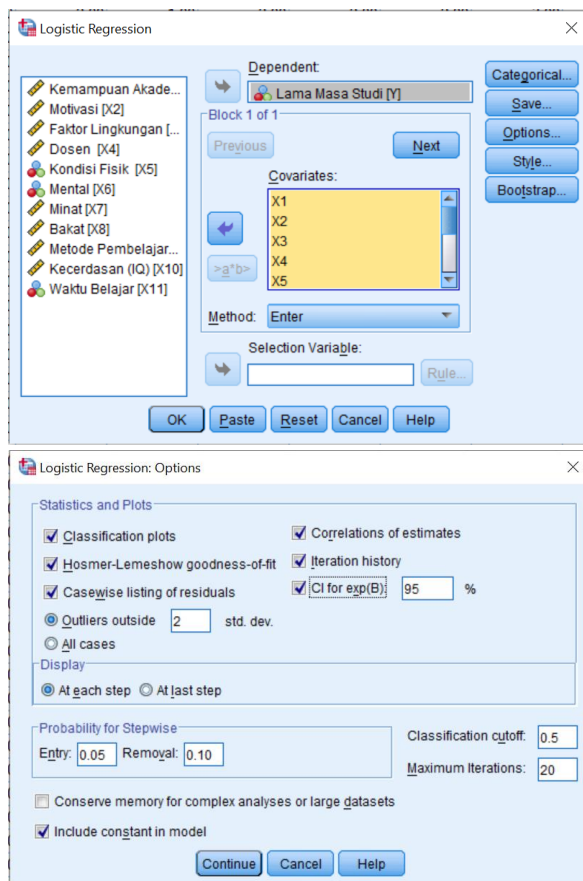


Figure 4. Logistic Regression dan Logistic Regression Option

5. Next SPSS 22.0 will display the calculation results of the Binary Logistics Regression Method. The results are as follows:

**Table 2. Hosmer and Lemeshow Test**

Step	Chi-square	df	Sig.
1	13.167	8	.106

If the value of Sig. in the table  $\geq 5\%$  then  $H_0$  is accepted and the hypothesized model fits the data. Table 2. Hosmer and Lemeshow Test show the value of Sig. of  $0.106 \geq 0.05$ , which means that the independent variable data (X1, X2, ..., X11) and the dependent variable can be analyzed by binary logistic regression test.

Table 3. Classification Table

Observed			Predicted		
			Study Period		Percentage Correct
			$\geq 4$ Years	$< 4$ Years	
Step 1	Study Period	$\geq 4$ Years	15	6	71.4
		$< 4$ Years	5	28	84.8
Overall Percentage					79.6



a. The cut value is .500

Table 3. Shows how much the results of the classification accuracy test for the binary logistic regression method are. Based on the classification table, 15 respondents are predicted to feel that the length of their undergraduate study period is more than or equal to 4 years a total of 21 respondents, so the classification accuracy is 71.4%. Meanwhile, as many as 28 respondents are predicted to feel that the length of their study period is less than 4 years from 34 respondents so the accuracy of the classification is 84.8%. This means that this binary logistic regression model can predict someone whose study period is less than 4 years where the person has a study period of fewer than 4 years or predicts someone whose study period is more than or equal to 4 years where the person is the length of the study period is more than or equal to 4 years is 79.6%.

Table 4. Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	52.427 <sup>a</sup>	.306	.415
a. Estimation terminated at iteration number 20 because maximum iterations have been reached. The final solution cannot be found.			

The coefficient of determination is the test result that shows how much the independent variable simultaneously or simultaneously affects the dependent variable. In table 4. To find out how much the independent variable affects the dependent variable simultaneously can be seen by paying attention to the value of Nagelkerke R Square. The Nagelkerke R Square value in the model summary table above is  $0.415 \times 100\% = 41.5\%$ . This means that simultaneously the independent variable (X) affects the dependent variable (Y) by 41.5%.

Table 5. The Best Model of Study Period

	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
							Lower	Upper
Step X1	.049	.706	.005	1	.944	1.051	.263	4.191
1 <sup>a</sup> X2	1.199	.854	1.974	1	.160	3.317	.623	17.672
X3	.906	.786	1.329	1	.249	2.474	.530	11.534
X4	1.117	.723	2.388	1	.122	3.057	.741	12.610
X6	-19.321	28044.404	.000	1	.999	.000	.000	.
X7	-.197	.737	.072	1	.789	.821	.194	3.479
X8	-.592	.707	.701	1	.402	.553	.139	2.211
X9	.755	.838	.813	1	.367	2.129	.412	11.000
X10	-1.834	.815	5.064	1	.024	.160	.032	.789
X11	.702	.793	.782	1	.376	2.018	.426	9.555

Constant	34.137	56088.808	.000	1	1.000	669096516805098.800		
----------	--------	-----------	------	---	-------	---------------------	--	--

a. Variable(s) entered on step 1: X1, X2, X3, X4, X6, X7, X8, X9, X10, X11.

The results of the analysis obtained a significant value of the Wald Test on the X10 variable of  $0.024 < 0.05$  so that the independent variable X10, namely intelligence (IQ) partially influences the dependent variable. This means that intelligence (IQ) has the most partial influence on the length of the undergraduate study period (S1). The value of Exp (B) on the X10 variable is 0.160, which means that high intelligence (IQ) affects the length of the undergraduate study period (S1) more than medium and low intelligence (IQ).

$$\pi(x) = \frac{\exp \exp (34.137 - 1.834 x_{10})}{(34.137 - 1.834 x_{10})}$$

### Conclusion

The binary logistic regression model of the length of the study period for Lecturers, Employees, and Students at the Universitas Pembangunan Panca Budi is as follows:

$$\pi(x) = \frac{\exp \exp (34.137 - 1.834 x_{10})}{(34.137 - 1.834 x_{10})}$$

From the results of the analysis, it is found that the percentage of classification accuracy in predicting the independent variable in influencing the dependent variable is 79.6%. partially Undergraduate study period (S1) compared to moderate and low Intelligence (IQ). And simultaneously all the factors that affect the length of the study period are academic ability, motivation, environment, lecturers, physical condition, mental, interests, talents, learning methods, intelligence (IQ), and study time by 41.5%, the rest influenced by other factors.

### Recommendation

### Acknowledgements

### References

- Aunurrahman. 2009. *Belajar dan Pembelajaran*. Bandung: Alfabeta.  
 Ginting, Cipta. 2005. *Kiat Belajar di Perguruan Tinggi*. Jakarta: Grasindo.  
 Hakim, Thursan. 2005. *Belajar Secara Efektif*. Jakara : Puspa Swara.  
 Hosmer, D.W. dan Lemeshow, S. (1989), *Applied Logistic Regression*, John Wiley & Sons, Inc., New York.  
 Mulyati, Cahyadi, Yuniar. 2019. Faktor-Faktor yang Mempengaruhi Penyelesaian Masa Studi Program Pascasarjana di Institut Pertanian Bogor (Thesis). Bogor : Program Pascasarjana, IPB University.

Pratiwi, Handoyo, Murti Nugraha. 2016. *Faktor-Faktor yang Mempengaruhi Masa Studi Mahasiswa Pendidikan Teknik Bangunan Universitas Negeri Jakarta*. Jurnal Pendidikan Teknik Sipil. Vol 5(2). 118 – 127.

Slameto. 2010. *Belajar dan Faktor-Faktor yang Mempengaruhinya*. Jakarta: PT Rineka Cipta