JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol XI No 3, Juni 2025, hlm. 525 – 532 ISSN 2550-0201 (Online)
DOI: https://doi.org/10.33330/jurteksi.v11i3.3893
Available online at https://jurnal.stmikroyal.ac.id/index.php/jurteksi

PREDICTION OF ON-TIME GRADUATION OF UNIVERSITAS ROYAL STUDENTS USING MULTIPLE LINEAR REGRESSION METHOD

Nurul Rahmadani[1*], Edi Kurniawan[1], Nurhasanah [2], Wahdan Damanik [1]
[1]Information System, Universitas Royal
[2]Management, Universitas Royal
email: cloudyrara@gmail.com

Abstract: On-time graduation is an important indicator in measuring the success of higher education and reflects the effectiveness of the academic process in higher education. Royal University, especially the Information Systems Study Program, still faces challenges in increasing the percentage of students who graduate on time. This study aims to identify factors that influence students' on-time graduation and build a prediction model using the multiple linear regression method. This method was chosen because it is able to analyze the simultaneous influence of several independent numeric variables on one dependent variable, making it suitable for studying the complex relationship between factors that influence student graduation. The independent variables analyzed in this study include GPA, parental income, and student part-time jobs with student graduation as the dependent variable. The results showed that parental income and part-time jobs had a significant positive effect on on-time graduation, while GPA had a negative effect. The model built had an R^2 value of 0.6153 and a standard error of 4.0653, indicating that the model was quite strong and accurate. These findings recommend Universitas Royal to strengthen the academic monitoring system and support working students, as well as design policies based on students' socio-economic conditions to increase the on-time graduation rate.

Keywords: multiple linear regression; on-time graduation; students.

Abstrak: Kelulusan tepat waktu merupakan indikator penting dalam mengukur keberhasilan pendidikan tinggi serta mencerminkan efektivitas proses akademik di perguruan tinggi. Universitas Royal, khususnya Program Studi Sistem Informasi, masih menghadapi tantangan dalam meningkatkan persentase mahasiswa yang lulus tepat waktu. Penelitian ini bertujuan untuk mengidentifikasi faktor-faktor yang memengaruhi kelulusan tepat waktu mahasiswa serta membangun model prediksi menggunakan metode regresi linear berganda. Metode ini dipilih karena mampu menganalisis pengaruh simultan beberapa variabel independen numerik terhadap satu variabel dependen, sehingga sesuai untuk mengkaji hubungan kompleks antar faktor yang memengaruhi kelulusan mahasiswa. Variabel independen yang dianalisis dalam penelitian ini meliputi IPK, penghasilan orangtua, dan pekerjaan sambilan mahasiswa dengan kelulusan mahasiswa sebagai variabel dependen. Hasil penelitian menunjukkan bahwa penghasilan orangtua dan pekerjaan sambilan berpengaruh positif signifikan terhadap kelulusan tepat waktu, sedangkan IPK justru memiliki pengaruh negatif. Model yang dibangun memiliki nilai R^2 sebesar 0,6153 dan standar error 4,0653, menandakan model cukup kuat dan akurat. Temuan ini merekomendasikan Universitas Royal untuk memperkuat sistem monitoring akademik dan mendukung mahasiswa yang bekerja, serta merancang kebijakan berbasis kondisi sosial-ekonomi mahasiswa guna meningkatkan angka kelulusan tepat waktu.

Kata kunci: kelulusan tepat waktu; mahasiswa; regresi linear berganda.

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol XI No 3, Juni 2025, hlm. 525 – 532 ISSN 2550-0201 (Online)
DOI: https://doi.org/10.33330/jurteksi.v11i3.3893
Available online at https://jurnal.stmikroyal.ac.id/index.php/jurteksi

INTRODUCTION

On-time student graduation is an important indicator that shows the success of higher education in Indonesia. This is reflected in the determination of the Ministry of Education, Culture, Research, and Technology regarding Key Performance Indicators (KPI). On-time graduation reflects the effectiveness of the academic process and has an impact on the quality of education produced by the institution [1]. Students are considered to have graduated on time if they have successfully completed all courses, completed their thesis, and achieved the university's academic standards [2]. Therefore, measuring the actual duration against the ideal duration is an important tool in assessing the success of an institution [3].

On-time graduation also has an impact on the personal and professional aspects of students [4]. In addition to being an academic achievement, on-time graduation shows commitment and discipline, which are highly valued in the world of work. Yatimah [5] highlighted that late graduation can affect the competitiveness of graduates in the job market. Fitria et al. [4] stated that the on-time graduation rate contributes to the good reputation of the university and the job prospects of its alumni.

Student academic success is closely correlated with the image of the institution in the eyes of the public. The rate of on-time graduation affects accreditation and attractiveness to prospective new students [6]. In addition, this success also opens up opportunities for strategic cooperation with various other institutions [7].

The success of Royal University in organizing higher education also depends heavily on the timely graduation of its students. Universitas Royal, located at Jalan Prof. H.M. Yamin, No. 173, Asahan Regency, North Sumatra Province, has three faculties and seven study programs. In this university environment, timely graduation is an indicator of the quality of education delivery and the effectiveness of the academic guidance system [1][7]. Given the large number of study programs available, this study focuses on the graduation of students from the Sistem Informasi Study Program.

However, many students experience delays in graduation due to academic and non-academic factors. Therefore, a deep understanding of the factors that influence on-time graduation is needed [8]. Hartati et al. [9] encourage the use of analytical methods to build predictive models that are useful in monitoring student academic performance.

The multiple linear regression method was chosen in this study because it is able to explore the relationship between independent variables (e.g. GPA, parental income, and part-time jobs) to the dependent variable (on-time graduation). The selection of the right independent variables is very important in the analysis and prediction process. The advantage of this method lies in its ability to analyze numerical data with linear relationships between variables [10], which makes it more suitable than methods such as Naive Bayes, especially when there is no dominant categorical classification [9].

Several previous studies have discussed the prediction of on-time graduation of students using various approaches and methods. Dengen et al. (2020) studied student graduation with a decision tree algorithm. The study used the attributes of gender, GPA, graduation predicate, and TOEFL score as variables to

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol XI No 3, Juni 2025, hlm. 525 – 532 ISSN 2550-0201 (Online)
DOI: https://doi.org/10.33330/jurteksi.v11i3.3893
Available online at https://jurnal.stmikroyal.ac.id/index.php/jurteksi

predict on-time graduation of students. With the help of the Weka application, this study achieved a prediction accuracy rate of 60%. Although effective in classification, this approach focuses more on forming a decision tree model and does not evaluate the influence of each variable quantitatively [2].

Meanwhile, Hartati et al. (2022) applied the logistic regression method to predict student graduation in the pass or fail category. Although their approach provides fairly accurate classification results, the use of logistic regression tends to limit the analysis to binary categories without evaluating the numerical influence measurably [7]. On the other hand, Yatimah (2023) focuses more on qualitative studies related to the influence of part-time work on student graduation. The single focus on the work variable makes the scope of the study limited and does not touch on other factors that may contribute [5].

Different from previous studies, this study presents a quantitative approach using multiple linear regression methods that are able to analyze the simultaneous influence of several independent variables, namely GPA, parental income, and part-time jobs on students' on-time graduation. The selection of this method is based on its ability to explain the linear relationship between numerical variables in more detail and accurately.

The novelty of this study lies in predictive modeling that combines academic and socio-economic factors in one integrated analysis, which is specifically applied to the local context of the Universitas Royal, Sistem Informasi Study Program. In addition, this study also provides practical contributions in the form of data-based policy recommendations that can be used by universities to increase students' on-time graduation rates more effectively and in a targeted manner.

METHOD

Multiple linear regression is one of the statistical techniques in data mining that is used to predict the value of the dependent variable (Y) based on more than one independent variable (X_1, X_2, ..., X_n). In addition, it is also to determine the pattern of the relationship between variable X and variable Y [11][12]. The multiple linear regression equation can be calculated using the formula (1):

$$Y' = b_0 + b_1X_1 + b_2X_2 + ... + b_nX_n \quad (1)$$

Description:
Y' = dependent variable (prediction value)
$X_1, X_2, ...X_n$ = independent variable
b_0 = coefficient (intercept)
$b_1, b_2, ...b_n$ = regression coefficient

To calculate b_0, b_1, b_2, ..., b_n, the least squares method is used with the following equation (2):

$$\begin{aligned}
b_0 n &+ b_1 \sum X_1 + b_2 \sum X_2 + \ldots + b_n &&= \sum Y \\
b_0 \sum X_1 &+ b_1 \sum X_1^2 + b_2 \sum X_1X_2 + \ldots + b_n \sum X_1X_n &&= \sum X_1Y \\
b_0 \sum X_2 &+ b_1 \sum X_1X_2 + b_2 \sum X_2^2 + \ldots + b_n \sum X_2X_n &&= \sum X_2Y \\
&\vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots && \quad (2) \\
b_0 \sum X_n &+ b_1 \sum X_1X_n + b_2 \sum X_2X_n + \ldots + b_n \sum X_n^2 &&= \sum X_nY
\end{aligned}$$

The next step is to determine the values of b_0, b_1, b_2, and b_3 based on the matrix, in the following formula equation (3):

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol XI No 3, Juni 2025, hlm. 525 – 532 ISSN 2550-0201 (Online)
DOI: https://doi.org/10.33330/jurteksi.v11i3.3893
Available online at https://jurnal.stmikroyal.ac.id/index.php/jurteksi

$$\hat{b}_0 = \frac{Det\ A_1}{Det\ A} \quad \hat{b}_2 = \frac{Det\ A_3}{Det\ A} \quad \hat{b}_1 = \frac{Det\ A_2}{Det\ A} \quad \hat{b}_3 = \frac{Det\ A_4}{Det\ A} \tag{3}$$

The coefficient of determination is a statistical value used to determine the relationship of influence between two variables. The coefficient of determination value shows the percentage of variation in the value of the dependent variable explained by the results of the regression equation [13]. The r^2 value ranges from $0 < r^2 < 1$.

To facilitate interpretation of the strength of the relationship between variables, it will be explained in the following table 1.

Table 1. Correlation Between Variables

r^2 value	Description
0	No correlation
>0 – 0,25	Very weak
>0,25 – 0,5	Quite strong
>0,5 – 0,75	Strong
>0,75 – 0,99	Very strong
1	Perfect correlation

To find out the value of r2, it can use the following formula equation (4).

$$r^2 = \frac{(b_1 \sum X_1 Y) + (b_2 \sum X_2 Y) + (b_3 \sum X_3 Y)}{\sum Y^2} \tag{4}$$

Standard error (SE) is used to measure how far the data is spread from the regression line. The smaller the SE value, the more accurate the model is in predicting. The standard error formula can be seen in the following equation (5).

$$Se = \sqrt{\frac{\sum (Y - Y')^2}{n-k}} \tag{5}$$

Description:
Se = Standard estimate difference
Y = actual value
Y' = predicted value
n = number of data
k = number of variables

RESULT AND DISCUSSION

In predicting students' on-time graduation, a system is needed that can assist in the process. In this study, the data used is student graduation data totaling 113 students. Where the independent variables (X) used in this study include the Average Grade Point Average (GPA), parental income, and part-time work of students, while the dependent variable (Y) is the time of student graduation.

However, this study only discusses the calculation process of the linear regression method calculation. This study analyzes the linear regression process to determine the optimal results and in the future will be implemented into a more integrated information system. For more details, can be seen in Table 2.

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol XI No 3, Juni 2025, hlm. 525 – 532 ISSN 2550-0201 (Online)
DOI: https://doi.org/10.33330/jurteksi.v11i3.3893
Available online at https://jurnal.stmikroyal.ac.id/index.php/jurteksi

Table 2. Student Graduation Data

No.	Students	GPA	Parents Income	Student Part Time Jobs	Graduation Time
1	Student-1	3,82	Medium	Not working	4 Years 3 Months
2	Student-2	3,72	Medium	Working	4 Years 5 Months
3	Student-3	3,62	Medium	Not working	4 Years 5 Months
4	Student-4	3,08	Low	Working	4 Years 6 Months
5	Student-5	3,14	Medium	Working	4 Years 7 Months
6	Student-6	3,67	High	Working	4 Years 7 Months
7	Student-7	3,85	Medium	Not working	3 Years 7 Months
8	Student-8	3,60	Medium	Not working	3 Years 7 Months
9	Student-9	3,00	High	Working	6 Years 7 Months
10	Student-10	3,56	Medium	Working	4 Years 7 Months
⋮	⋮	⋮	⋮	⋮	⋮
113	Student-113	3,63	Low	Not working	3 Years 8 Months

The student graduation data is first transformed into numeric form, so that it can be processed into the Multiple Linear Regression method, as in Table 3 below.

Table 3. Student Graduation Data Transformation

No.	Students	GPA	Parents Income	Student Part Time Jobs	Graduation Time
1	Student-1	3,82	2	1	51
2	Student-2	3,72	2	2	53
3	Student-3	3,62	2	1	53
4	Student-4	3,08	1	2	54
5	Student-5	3,14	2	2	55
6	Student-6	3,67	3	2	55
7	Student-7	3,85	2	1	43
8	Student-8	3,60	2	1	43
9	Student-9	3,00	3	2	79
10	Student-10	3,56	2	2	55
⋮	⋮	⋮	⋮	⋮	⋮
113	Student-113	3,63	1	1	44

After completing the data transformation, the next step is to find the equation $Y' = b_0 + b_1X_1 + b_2X_2 + ... + b_nX_n$. However, before that, the researcher must first find the values of b_0, b_1, b_2, and b_3, Here is the process:

$$b_0 = \frac{\text{Det } A_1}{\text{Det } A} = \frac{27446141,5}{354194,80} = 77,489$$

$$b_1 = \frac{\text{Det } A_2}{\text{Det } A} = \frac{-4280560,61}{354194,80} = -12,085$$

$$b_2 = \frac{\text{Det } A_3}{\text{Det } A} = \frac{1080911,59}{354194,80} = 3,052$$

$$b_3 = \frac{\text{Det } A_4}{\text{Det } A} = \frac{1979690,31}{354194,80} = 5,589$$

Based on these results, the Multiple Linear Regression equation for predicting student graduation is:

$$Y' = 77{,}489 + (-12{,}085)X_1 + 3{,}052X_2 + 5{,}589X_3$$

The regression equation obtained, it can be seen the influence of each independent variable on the dependent variable, namely student graduation. The results of the study showed that GPA (variable X_1) has a negative influence on graduation, with a coefficient of -12,085. Meanwhile, parental income (X_2) and part-time work (X_3) have a significant positive influence, each with a coefficient value of 3,052 and 5,589.

Among the three variables, part-time work has the greatest influence on student graduation at Royal University. Students who have part-time jobs tend to experience delays in graduation because their time and focus are divided, potentially hindering the completion of studies on time.

The next step is to find the coefficient of determination (r^2) value to determine the correlation between variables.

$$r^2 = \frac{(b_1 \sum X_1 Y) + (b_2 \sum X_2 Y) + (b_3 \sum X_3 Y)}{\sum Y^2}$$

$$r^2 = \frac{(-12{,}085 * -104{,}25) + (3{,}052 * 122{,}73) + (5{,}589 * 223{,}05)}{4682{,}50}$$

$$r^2 = \frac{2881{,}114}{4682{,}50} = 0{,}6153$$

Based on the calculation, the coefficient of determination (r2) was obtained as 0.6153, which means Strong Correlation. In addition, it can be concluded that the influence of GPA (X_1), Parental Income (X_2), and Student Part-time Jobs (X_3) on students' on-time graduation (Y) is 61.53%. The remaining 38.47% is caused by other factors not included in the model.

The next step is to find the Standard Error (SE) value which functions to measure how far the data is spread from the regression line. The smaller the SE value, the more accurate the model is in predicting. So, to find out the standard error value in this study, it can be seen in the following calculation:

$$Se = \sqrt{\frac{\sum (Y - Y')^2}{n - k}}$$

$$Se = \sqrt{\frac{1801{,}382}{113 - 4}}$$

$$Se = \sqrt{16{,}5624}$$

$$Se = 4{,}0653$$

So, the standard Error result is 4.0653, which means that the error rate in the calculation using this model has a small error rate.

CONCLUSION

This study successfully built a prediction model for on-time graduation of Universitas Royal students, especially the Information Systems Study Program, using the multiple linear regression method. This model was tested on 113 student data and produced a coefficient of determination (R^2) value of 0.6153 which indicates a strong correlation, and a standard error level of 4.0653 which indicates a relatively small prediction error. These results prove that the model has the potential to be used in monitoring and eval-

uating student graduation in a more measurable way. Therefore, Universitas Royal is advised to implement this model periodically in the academic system as a prediction and decision-making tool. In addition, the prediction results can be used by academic advisors and academics to develop more personalized guidance strategies. Considering that parental income and part-time jobs have been shown to have a significant effect on graduation, additional support in the form of scholarships or flexible work programs is needed. In the future, the development of this data-based predictive system can also be integrated into the academic information system to provide automatic warnings to students who have the potential to experience delays in graduation, in order to support the achievement of on-time graduation targets as a whole.

BIBLIOGRAPHY

[1] M. J. Kurniawan and G. Wang, "Gamifikasi Sistem Pemantauan Kinerja Mahasiswa Dalam Mencapai Kelulusan Tepat Waktu," J. Teknol. Sist. Inf., vol. 5, no. 1, pp. 105–118, 2024, doi: 10.35957/jtsi.v5i1.7785.

[2] C. N. Dengen, K. Kusrini, and E. T. Luthfi, "Implementasi Decision Tree Untuk Prediksi Kelulusan Mahasiswa Tepat Waktu," Sisfotenika, vol. 10, no. 1, p. 1, 2020, doi: 10.30700/jst.v10i1.484.

[3] F. F. Anwar, A. I. Jaya, and M. Abu, "Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Decision Tree Dengan Penerapan Algoritma C4.5," J. Ilm. Mat. Dan Terap., vol. 19, no. 1, pp. 19–28, 2022, doi:10.22487/2540766x.2022.v19.i1.15880.

[4] N. Fitria, A. C. H. CHAIRY, and A. C. F. NAHDIYAH, "Mengurai Kompleksitas Budaya Organisasi Dalam Konteks Manajemen Pendidikan Pada Lembaga Pendidikan Tinggi," J. Kepengawasan Supervisi Dan Manajerial, vol. 1, no. 4, pp. 136–142, 2024, doi: 10.61116/jksm.v1i4.257.

[5] M. N. Yatimah, "Implementasi Data Mining Untuk Prediksi Kelulusan Tepat Waktu Mahasiswa STIMIK ESQ Menggunakan Decision Tree C4.5," Jumanji (Jurnal Masy. Inform. Unjani), vol. 5, no. 2, p. 89, 2021, doi: 10.26874/jumanji.v5i2.95.

[6] N. Khasanah, A. Salim, N. Afni, R. Komarudin, and Y. I. Maulana, "Prediksi Kelulusan Mahasiswa Dengan Metode Naive Bayes," Technol. J. Ilm., vol. 13, no. 3, p. 207, 2022, doi: 10.31602/tji.v13i3.7312.

[7] S. Hartati and H. A. SAN, "Algoritma Naive Bayes Untuk Prediksi Kelulusan Mahasiswa," J. Cakrawala Inf., vol. 2, no. 2, pp. 42–50, 2022, doi: 10.54066/jci.v2i2.234.

[8] D. Safitri, S. S. Hilabi, and F. Nurapriani, "Analisis Penggunaan Algoritma Klasifikasi Dalam Prediksi Kelulusan Menggunakan Orange Data Mining," Rabit J. Teknol. Dan Sist. Inf. Univrab, vol. 8, no. 1, pp. 75–81, 2023, doi: 10.36341/rabit.v8i1.3009.

[9] S. Hartati, N. A. Ramdhan, and H. A. SAN, "Prediksi Kelulusan Mahasiswa Dengan Naïve Bayes

Dan Feature Selection Information Gain," J. Ilm. Intech Inf. Technol. J. Umus, vol. 4, no. 02, pp. 223–234, 2022, doi: 10.46772/intech.v4i02.889.

[10] M. Lase, D. Saripurna, and V. W. Sari, "Estimasi Penjualan Ice Cream Walls Menggunakan Metode Regresi Linear Berganda," J. Sist. Inf. Triguna Dharma (Jursi Tgd), vol. 1, no. 5, p. 625, 2022, doi: 10.53513/jursi.v1i5.5146.

[11] S. A. Sakroni, "Sistem Peramalan Jumlah Penumpang Di Bandarasultan Syarif Kasim II Menggunakan Metode Regresi Linier Berganda Sistem Peramalan Jumlah Penumpang Di Bandara Sultan Syarif Kasim Ii Menggunakan Metode Regresi Linier Berganda," 2022.

[12] T. Azhara, N. Rahmadani, and A. Nata, "Application of Multiple Linear Regression Estimating the Population of Asahan Regency," JURTEKSI (Jurnal Teknol. dan Sist. Informasi), vol. 10, no. 1, pp. 1–8, 2023, doi: 10.33330/jurteksi.v10i1.2466.

[13] V. N. Yulianti and P. Sembiring, "Penerapan Metode Backward untuk Menentukan Persamaan Regresi Linier Berganda pada Dugaan Tindak Pidana di Kota Binjai," FARABI J. Mat. dan Pendidik. Mat., vol. 6, no. 1, pp. 1–9, 2023.