Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 1, January 2022
DOI : https://doi.org/10.33395/sinkron.v7i1.11261
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Student Graduation Predictions using Comparison of C5.0 Algorithm with Linear Regression

Fevi Ariska[1)*], Volvo Sihombing[2)], Irmayani[3)]
[1)2)3)]Labuhanbatu University, North Sumatra, Indonesia
[1)] feviariska23@gmail.com , [2)] volvolumbantoruan@gmail.com , [3)] irmayantiritonga2@gmail.com

Submitted : Jan 15, 2022 | **Accepted** : Feb 1, 2022 | **Published** : Feb 3, 2022

Abstract : Technological advances supported by human knowledge have a very good influence on data and information storage technology, including in predicting student graduation (Graduation Prediction) on time, by applying several existing algorithms. In this study, researchers used the C5.0 Algorithm and Linear Regression. The concept of the research is to compare two algorithms, namely C5.0 and Linear Regression to the case of graduating students on time. Based on the length of study, students who graduated correctly amounted to 651 (91%) with a male gender of 427 students and a female gender of 224 students while those who did not pass (late) correctly amounted to 64 (9%) with a male gender totaling 53 students and female gender totaling 11 students from 2017-2020. Comparison results The R2 score from the C5.0 algorithm reached 96.85% (training) and 93.

Keywords: C5.0 Algorithm, Graduation Prediction, Linear Regression

INTRODUCTION

With the development of technological progress and supported by human knowledge, it has a very good influence on the technology of data and information deviation. With the development of technology in all fields so that obtain a lot of information in the form of data in the form of economic industry, and others. The information obtained is very much but the data is still a lot that has not been used properly.

The data is processed by analyzing data called data mining, the concept of this data is the process of finding selected data patterns, using methods.Knowledge Discovery in Database (KDD) is a process summary important information from large databases. Data mining divides the data model obtained in the data set, so that the data can be optimized for use in everyday life (DS Panggabean et al., 2020)

. Big data is large data collected in the form of very valuable data sets and can be used in the analysis process, where the results obtained can be in the form of knowledge or information so that they can be used for now as a reference for comparison with previous data. This process does not only apply in the business world but can be applied to agencies such as universities that have a lot of data (E. Setiawan et al., 2019).

at college tall, there are a lot of student data and data on the number of graduations each year that is able to produce information. One of the achievement parameters of a higher education institution is the presentation of student graduations who pass exactly in accordance with the regulations of the higher education national accreditation body number 3 in 2019.

One of the universities that has a lot of data is Prima Indonesia University where one of them is the faculty technology and computer science. The faculty of technology and computer science summarizes student data in a student academic information system that has a database and manages it as needed. Based on the data obtained, it is found that the faculty of technology and computer science, had 174 students in 2017, and 156 people graduated on

* Corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 1, January 2022
DOI : https://doi.org/10.33395/sinkron.v7i1.11261
e-ISSN : 2541-2019
p-ISSN : 2541-044X

time, so that the graduation percentage was 89%. In 2018 there were 268 students, 240 students graduated on time, and a passing percentage of 89% was obtained, while in 2020 there were 273 students, and 254 students graduated on time, so the student graduation percentage was 93%.

In accordance with the results of data collection by researchers and the results of existing data, each new student admission is accepted the more increases, but not all students can graduate on time which will result in data accumulation so that incoming data and outgoing data are out of sync. This results in an increase in the amount of data and academic data for all faculties from all students who are still registered every year (MM Baharuddin et al., 2019).

There are also studies that have been carried out to predict student graduation on time at other universities by applying existing algorithms. One of them is a decision tree. There are several cases that have been carried out as prediction of students' timely graduation using the naive bayes algorithm (AAMurtopo., 2016). There is also the application of classification techniques to predict student graduation by applying the K-Nearest Neighbor algorithm (AY Saputra et al., 2018), there are also researchers comparing the C4.5 and Naive Bayes algorithms with the case of students graduating on time (M. Kamil et al. .,2020). The analysis predicts in the case of students dropping out by applying the decision tree method with the ID3 and C4.5 Algorithms (LR Haidar et al., 2020), there are also cases in predicting student graduation using the decision tree method using the C.45 algorithm. The existence of existing research results and problems within the scope of the faculty regarding the number of students, the research team is interested in conducting a study, namely predicting student graduation using the comparison of the C5.0 algorithm with linear regression which will be used to predict the graduation of students who will be graduating. tested on students of the faculty of information technology and computer science. In accordance with the results of the study, it is hoped that it can be a benchmark for improving the quality of students at Universita Prima Indonesia (Mashlahah., 2013).

LITERATURE REVIEW

Data Mining

according to(Fahmi & Sianturi, 2019)The human need for data and information cannot be denied. In fact, now through the world of technology, the flow of information can circulate quickly and easily. Data needs to be organized and controlled into information to make it easier to understand. Processing of data into information needs to be done carefully so that the resulting information has good quality. Along with this, data mining algorithms are also developing on large databases. Data mining is a technology that is useful for extracting knowledge or what is known as information from a collection of data, so that the results can be used for decision making.

Decision Tree

Decision tree or decision tree is a tree that is used as a reasoning procedure to get answers to the problems entered. The tree that is formed is not always a binary tree. If all the features in the data set use 2 kinds of categorical values, then the tree form obtained is in the form of a binary tree. If the feature contains more than 2 kinds of categorical values or uses a numeric type, the tree form obtained is usually not a binary tree(Sianturi et al., 2018) According to Prasetyo (Sianturi, 2018)Decision tree has three classical approaches: 1. Classification tree, used to make predictions when there is new data whose class label is not known. This approach is the most widely used. 2. Regression tree, when the predicted results are considered as real values that are likely to be obtained. For example the case of rising house prices, inflation predictions every year, and so on. 3. CART (or C&RT), when classification and regression problems are used together.

C5.0 . Algorithm

The C5.0 algorithm is one of the data mining algorithms which is especially applied to the decision tree algorithm. The C5.0 algorithm is a refinement of the previous algorithm formed by Ross Quinlan in 1987, namely ID3 and C4.5. In this algorithm, the selection of attributes is processed using the gain ratio. This algorithm produces a tree with a variable number of branches per node(Mardi, 2017)

The C5.0 algorithm produces a tree with varying number of branches per node. This algorithm treats continuous variables the same as CART, but for categorical variables the C5.0 algorithm treats the values of categorical variables as splitters. The subset sample obtained from the formed branching will be broken down again afterwards. The process

* Corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 1, January 2022
DOI : https://doi.org/10.33395/sinkron.v7i1.11261
e-ISSN : 2541-2019
p-ISSN : 2541-044X

will continue until the subset samples can no longer be divided. In the end, the sample subset that does not have a large contribution to the model will be rejected(Elisa, 2017).

The working steps of tree creation in the C5.0 algorithm are similar to tree creation in the C4.5 algorithm. The similarities include the calculation of entropy and gain. If the C4.5 algorithm stops until the gain calculation, then the C5.0 algorithm will continue by calculating the gain ratio using the existing gain and entropy.

The formula for finding the entropy value is as follows:

$$Entropy = -\sum_{j=1}^{k} p_j log_2 p_j \quad (1)$$

With S: Set of cases ; k: number of classes in variable A ; the proportion of Sj and Sp_j:
Furthermore, to find the gain value, the following equation is used:

$$Gain\ (S, A) = Entropy(S) - \sum_{i=1}^{m} \frac{|S_i|}{|S|}\ x\ Entropy\ (S_i) \quad (2)$$

After obtaining the entropy and gain values, the next step is to calculate the gain ratio value. The basic formula for calculating the gain ratio is as follows:

$$Gain\ Ratio = \frac{Gain\ (S,A)}{\sum_{i=1}^{m} Entropy\ (S_i)} \quad (3)$$

The process is repeated for each branch until all classes in the branch have their own class.

3. METHOD STUDY

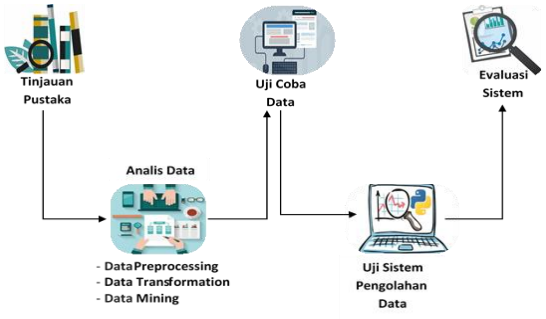

Fig.1 Research Methodology Flow

The initial stage of the researcher collects the data needed for the data mining process. By using from the faculty of technology and computer science. The Student/Alumni data taken has a total of 715 rows and 9 attributes/columns from 2017-2020. The data from this selection will be used in processing the data.

The initial stage carried out before analyzing is the data processing stage, it is necessary to carry out an analysis of the data that has been taken to be able to check the clarity and completeness of the research data. The purpose of data analysis is to determine the quality of the data to be tested and to draw hypotheses related to the data to be tested. There are several steps in analyzing the data, namely: (a) Preprocessing data, at this stage is the stage for data cleaning (data cleaning). The data cleaning carried out in this study is to remove data attributes that are not needed in the prediction process later and to delete data that is null or N/A. (b) data transformation, the next stage is the data transformation stage. At this stage, the data normalization process will be carried out. The purpose of this normalization is so that the data, which are especially numeric, are in the range 0 to 1 so that the data distribution is not too far away. Meanwhile, categorical data will be converted to numeric so that it can be matched in the next data mining process. (c) data mining, which is the process of looking for patterns that have been selected and transformed according to the C5.0 algorithm and linear regression, which later this algorithm has a comparative value against predictions of timely graduation for information technology and computer science students in 2017-2020 based on the total number registered students.

* Corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 1, January 2022
DOI : https://doi.org/10.33395/sinkron.v7i1.11261
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Python Trial is the implementation stage of analysis and prediction of Student Graduation of the Faculty of Information Technology and Science Computerat Prima University where a program written in Python version 3.7.12 will be tested on the Google Colab platform. Python is referred to as a programming language that is able to combine skills and abilities by using grammar through clear code, to give orders to the programs created [8].

System testingisthe execution stage of the software system to determine whether the implementation of Data Mining matches the system specifications. Testing this system is in the form of verifying whether the program that has been made can be run according to the specifications and designs that have been made.

This evaluation stage is the last stage carried out to produce the capacity and effectiveness of the model that is able to meet the objectives and can provide solutions when solving the problems in this research, as well as making decisions regarding the use of the results from Data Mining.

RESULTS

Data Analysis

Student Graduation Dataset

The researcher uses two methods, namely the C5.0 algorithm and linear regression to predict graduation of students from the faculty of technology and computer science which is implemented in a Python-based program. The data used is student graduation data in 2017-2020, it can be seen in Table 1.

Table 1
Student graduation sample data

NIM	Name	Study Program	No. Certificate	GPA	JK	Study Length
123303030078	Subanto	IT	294/S1-TI/FTIK/UNPRI/IV/2017	3.18	LK	Late
123303030084	Andre	IT	296/S1-TI/FTIK/UNPRI/IV/2017	2.67	LK	Late
123303030008	Novendy	IT	297/S1-TI/FTIK/UNPRI/IV/2017	3.49	LK	Appropriate
153303040241	Desi Gulo	SI	654/S1-SI/FTIK/UNPRI/VI/2020	2.47	PR	Late
163303030408	Cindy	SI	779/S1-SI/FTIK/UNPRI/IX/2020	3.70	PR	Appropriate
163303010406	Vanessa	TIND	68/S1-TIND/FTIK/UNPRI/IX/2020	3.82	PR	Appropriate
163303010407	Wilson	TIND	69/S1-TIND/FTIK/UNPRI/IX/2020	3.68	LK	Appropriate
163303020405	Wandy	TE	39/S1-TE/FTIK/UNPRI/IX/2020	3.70	LK	Appropriate
153303030444	Isnaini	SI	771/S1-SI/FTIK/UNPRI/IX/2020	3.34	PR	Late
163303030431	Steven	SI	749/S1-SI/FTIK/UNPRI/IX/2020	3.51	LK	Appropriate

This graduation data collection beginsfromgraduation year 2017 to 2020. Based on data received by researchers from the faculty of technology and computer science, the data can be collected into a data set under study. In this dataset there are several attributes used, namely NIM, Name, Study Program, No. Alumni, Diploma Number, GPA, Gender, Length of Study.

Analysis of Student Graduation

In this section will discuss about the analysis of student graduation by calculating the data that has been obtained from the results of research. The data that has been obtained will be analyzed and visualized from several variables so that it can be used as useful information.

Analysis of the Number of Students by Academic Year

Here the analysis is carried out on the number of students based on each study program (Prodi) based on the academic year, according to the pictureunder the number of students in the Informatics Engineering Study Program (61), Information Systems (91), Industrial Engineering (17), and Electrical Engineering (5) in 2017. In 2018 the Informatics Engineering Study Program (134), Information Systems (105), Industrial Engineering (19), and Electrical Engineering (10), while in 2020 the Informatics Engineering Study Program (109), Information Systems (118), Industrial Engineering (30), and Electrical Engineering (15).

* Corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 1, January 2022
DOI : https://doi.org/10.33395/sinkron.v7i1.11261
e-ISSN : 2541-2019
p-ISSN : 2541-044X

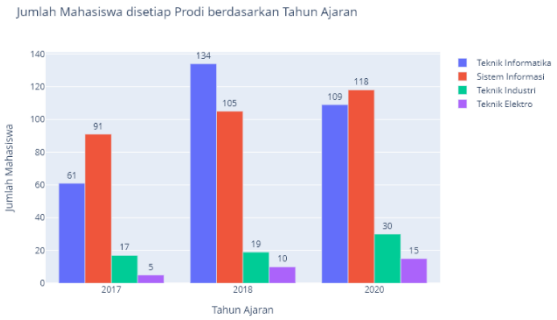

Fig.2 Analysis of Number of Students by Academic Year

Analysis of the Number of Students by Gender

At this stage the analysis donebased on the sex of students at the Faculty of Technology and Information, Prima Indonesia University. In Table.2. the number of students based on male gender totaled 480 students while the number of female students amounted to 235 students.

Table 2
Number of Students by Gender

Gender (JK)	Number of Students
Male (L)	480
Female (F)	235

From the results of calculations obtained based on table 2 above obtained the value of each gender, based on the data made the image below which is a visualization or display of the number of students based on gender with a view in the form of Percent (%), i.e. Men 67% while Women are 33%.

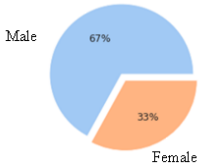

Fig.3 Visualization of Number of Students by Gender

Analysis of Student GPA by Study Program

Based on the calculations carried out by the results of the analysis of the calculations carried out, the study of this analysis was carried out on the student GPA (Cumulative Achievement Index) which is based on the student study program for the period 2017-2020. Here is an image of the results of the calculation analysis carried out.

* Corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 1, January 2022
DOI : https://doi.org/10.33395/sinkron.v7i1.11261
e-ISSN : 2541-2019
p-ISSN : 2541-044X

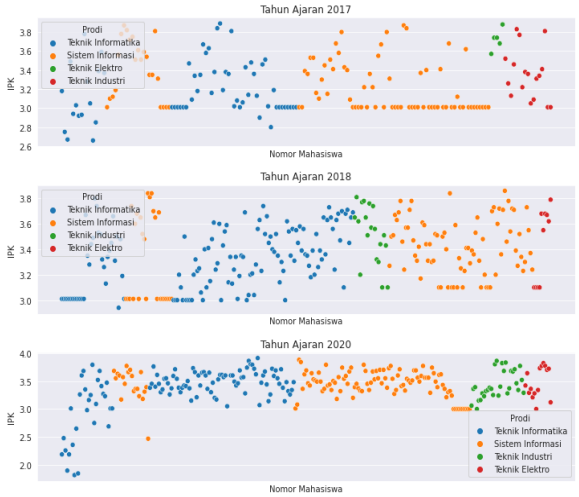

Fig.4 Analysis of Student GPA by Study Program

Analysis of Student Graduation Based on Length of Study

At this analysis stage, it will be indisplay a visualization of the graduation of the number of students based on the length of study for each study program. As in the picture below, it can be seen that the number of students who graduated correctly and did not graduate (late) in 2017, 2018 and 2020.

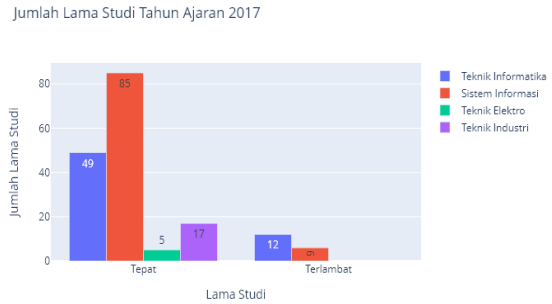

Fig.5 Analysis of Student Graduation Based on Study Length in 2017

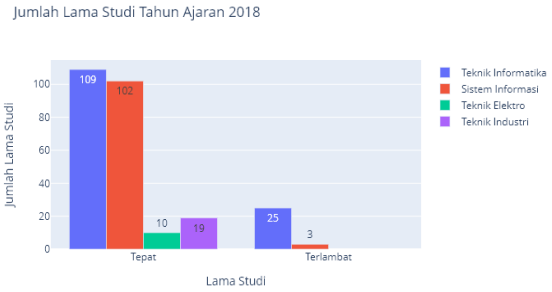

Fig.6 Analysis of Student Graduation Based on Length of Study in 2018

* Corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 1, January 2022
DOI : https://doi.org/10.33395/sinkron.v7i1.11261
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Fig.7 Analysis of Student Graduation Based on Length of Study in 2020

From the results of the visualization above, it can be concluded that the total number of students who graduated correctly were 651 (91%) with a male gender of 427 students and a female gender of 224 students while those who did not graduate correctly (late) amounted to 64 (9%) with there are 53 male students and 11 female students. As in the image that will appear.

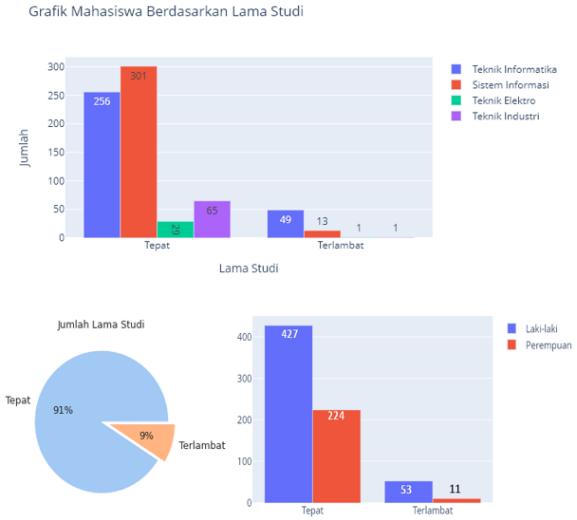

Fig.8 Analysis of the Overall Number of Graduates Based on Study Length in 2017-2020

3.1.7 Transformation of data on dataset

Where is it said that definition of data transformation is the process of replacing data into an appropriate form for processing in data mining, because sometimes the data used in data mining processing has a form that is not yet suitable so it cannot be used directly. With this system it is necessary to change its shape. The results of the change in the format formed in the table that has been grouped according to the predicted target, namely student graduation, in the data transformation stage requires several variables, including predictor variables: NIM, Name, Study Program, No. Diploma, GPA, Gender, Length of Study, Academic Year. Criteria Variable: “Right” or “Late”

* Corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 1, January 2022
DOI : https://doi.org/10.33395/sinkron.v7i1.11261
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Table 3
Dataset Transform

NIM	Name	Study Program	No. Certificate	GPA	JK	Study Length	School year
7	597	3	54	3.18	0	1	2017
8	31	3	56	2.67	0	1	2017
45	487	3	57	3.49	0	0	2017
446	152	0	491	2.47	1	1	2020
557	115	0	661	3.70	1	0	2020
467	645	2	519	3.82	1	0	2020
468	673	2	530	3.68	0	0	2020
441	539	1	143	3.34	0	1	2020
443	313	0	653	3.34	1	1	2020
572	593	0	677	3.51	0	0	2020

3.2 Trial Using the C5.0 . Algorithm

Every algorithm that has beentestedhas a significant value and has the characteristics of each algorithm. The C5.0 algorithm is an algorithm improvement from ID3 and C4.5. Where the C5.0 algorithm is said to be able to provide excellent data accuracy, where the error rate is low in cases that are not visible. And also can automatically delete data that is not important. In this algorithm, the selection of attributes is processed using the gain ratio [9]. If, the calculation of C4.5 is only to calculate the gain, while the C5.0 algorithm will continue the calculation of the gain ratio with the gain and entropy that have been determined.

The formula for finding the Entropy and Gain values is as follows:

$$\text{Entropy}\ (S) = -\sum_{i=1}^{n} p_i\ \log_2 p_i \quad (4)$$

$$\text{Gain}\ (S, A) = \text{Entropy}\ (S) - \sum_{i=1}^{m} \frac{|S_i|}{|S|}\ x\ \text{Entropy}(S_i) \quad (5)$$

Where is Entropy: S: Set case, n: number of partitions S, pi : Proportion of Si and S. and Gain : S : Set of cases, Si : Set of cases in the i-th category, A : Variable, m : Number of categories in variable A, Si: Number of cases in category I, S: Number of cases in S

From the experiment above, the results of the Student Graduation Prediction trial using the C5.0 Algorithm are as follows:

Table 4
C5.0 Algorithm Trial

actual	Predicted
0	1
0	0
0	0
0	1
1	1
0	0
0	0
0	0

* Corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 1, January 2022
DOI : https://doi.org/10.33395/sinkron.v7i1.11261
e-ISSN : 2541-2019
p-ISSN : 2541-044X

1	1
0	0

Testing Using Linear Regression

Linear Regression is an algorithm that involves the relationship between one dependent variable combined with one independent variable. Linear Regression serves to test the extent of the causal relationship between the factor variable (X) and the resultant variable (Y). Linear regression with one predictor variable is called Simple Regression, while for more than one predictor variable it is called Linear Regression [10].

In this study, the algorithm used is Linear Regression. For how Multiple Linear Regression works itself by predicting the effect of two or more predictor variables X1,....,Xn on one Y variable in order to prove the truth of the relationship between two or more input variables X1,...,Xn with one variable Y [1].

$$y = w_0 + w_1 x_1 + w_2 x_2 + \cdots + w_n x_n$$

Where is Linear Regression: y :Variable Criteria, w_0 : Bias, w_(i) : Coefficient, x_(i) : Predictor Variable, n : Total Data.

From the trial above, the results of the Prediction of Student Graduation were obtained using the Linear Regression Algorithm. The following are the results of trials using the Linear Regression Algorithm.

Table 5
Test Linear Regression Algorithm

actual	Predicted
0	-0.054588
0	0.361210
0	-0.136014
0	0.081159
1	0.421531
0	0.124975
0	-0.038992
0	0.081488
1	0.357250
0	0.131369

Evaluation of the Value of the C5.0 Algorithm with Linear Regression

The next step is the evaluation stage. This stage is the stage where the results of the trial are evaluated using an R2 Score. The value of R2 is used as a quantity to estimate the percentage variance of each algorithm. To get the evaluation value used the following formula:

$$\text{Accuracy} = \frac{TP+FP}{TP+FP+TN+FN} \qquad (6)$$

Where Accuracy: TP: True Positive, FP: False Positive, TN: True Negative, FN: False Negative.

From the evaluation results above, the results of the R2 Score comparison between the C5.0 Algorithm and Linear Regression can be made.

Table 6
Comparison of Training and Testing Accuracy Results on the C5.0 Algorithm with Linear Regression

Model	Training Accuracy (%)	Testing Accuracy (%)

* Corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 1, January 2022
DOI : https://doi.org/10.33395/sinkron.v7i1.11261
e-ISSN : 2541-2019
p-ISSN : 2541-044X

C5.0 . Algorithm	96.85	93.72
Linear Regression	33.31	40.30

DISCUSSION

From the results of research and testing trying data, researchers wrote that there are several conclusions in this study, namely from the results of calculations conducted based on old research, the number of students who graduated correctly was 651 (91%) with male sex being 427 students and female gender (P) to 224 students while those who did not graduate (late) were 64 (9%) with the gender of 53 male students (M) and 11 female students. from 2017-2020, as well as the R2 score comparison results from the C5.0 algorithm reached 96.85% (training) and 93.72% (testing) while the R2 score from Regression reached 33.31% (training) and 40.30% (testing).

CONCLUSION

Based on research conducted in the previous chapters, a conclusion was obtained that with the results of the comparison of the two algorithms between linear regression and C5.0 algorithm the value of higher percentage is with the calculation of the C5.0 algorithm, as a suggestion for researchers who will improve this research compared to several other methods related to prediction, to find out the extent of the accuracy of a method, and also use more data that will be used to get more perfect results.

REFERENCES

Baharuddin, M. M., Azis, H., & Hasanuddin, T. (2019). Analisis Performa Metode K-Nearest Neighbor untuk Identifikasi Jenis Kaca. *ILKOM Jurnal Ilmiah*, *11*(3), 269-274.

Elisa, E. (2017). Analysis and Application of the C4. 5 Algorithm in Data Mining to Identify Factors that Cause PT. Arupadhatu Adisesanti Construction Accidents. *J. Online Inform*, *2*(1), 36

Fahmi, M., & Sianturi, F. A. (2019). ANALYSIS OF APRIORIC ALGORITHM ON CONSUMER ORDERING AT CAFÉ THE L. COFFE COFFEE. *SAINTEK (Journal of Science and Technology., vol. 1, no. 1, pp. 52–57.*

Haidar, L. R., Sediyono, E., & Iriani, A. (2020). ANALISA PREDIKSI MAHASISWA DROP OUT MENGGUNAKAN METODE DECISION TREE DENGAN ALGORITMA ID3 dan C4. 5. *Jurnal Transformatika*, *17*(2), 97-106.

Herwanto, H. W., Widiyaningtyas, T., & Indriana, P. (2019). Penerapan Algoritme Linear Regression untuk Prediksi Hasil Panen Tanaman Padi. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi (JNTETI)*, *8*(4), 364-370.

Kamil, M., & Cholil, W. (2020). Analisis Perbandingan Algoritma C4. 5 dan Naive Bayes pada Lulusan Tepat Waktu Mahasiswa di Universitas Islam Negeri Raden Fatah Palembang. *Jurnal Informatika*, *7*(2), 97-106.

KP Wirdhaningsih, M. Ratnawati, Dian Eka, UB Malang, D. Mining, and D. Tree, Application of Decision Tree C5.0 Algorithm for Forex Forecasting, 2013th ed. Malang: Brawijaya University Malang, 2013.

Mardi, Y. (2017). *Data Mining :Classification Using the C4.5 Algorithm*. Journal of Informatics Education

Mashlahah, S., Yaqin, M. A., & Faisal, M. (2013). Prediction of Students Graduation Using Decision Tree Method with the Implementation of Algorithm C4. 5. *IEESE International Journal of Science and Technology*, *2*(2), 1.

Murtopo, A. A. (2016). Prediksi Kelulusan Tepat Waktu Mahasiswa STMIK YMI Tegal Menggunakan Algoritma Naïve Bayes. *CSRID (Computer Science Research and Its Development Journal)*, *7*(3), 145-154.

Panggabean, D. S. O., Buulolo, E., & Silalahi, N. (2020). Penerapan Data Mining Untuk Memprediksi Pemesanan Bibit Pohon Dengan Regresi Linear Berganda. *JURIKOM (Jurnal Riset Komputer)*, *7*(1), 56-62.

Saputra, A. Y., & Primadasa, Y. (2018). Penerapan Teknik Klasifikasi Untuk Prediksi Kelulusan Mahasiswa Menggunakan Algoritma K-Nearest Neighbor. *Techno. Com*, *17*(4), 395-403.

Setiawan, E., Antoni, D., & Mirza, A. H. (2019). Analisis Penerimaan sistem ujian online Berbayar Dengan Menggunakan METODE technology acceptance model (TAM) Dan WEBQUAL. *Jurnal Bina Komputer*, *1*(1), 61–72. https://doi.org/10.33557/binakomputer.v1i1.155

* Corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 1, January 2022
DOI : https://doi.org/10.33395/sinkron.v7i1.11261
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Sianturi, F. A. (2018). Decision Tree Analysis in Student Data Processing. *MEANS (Media Inf. Anal. And Sist., vol. 3, no. 2, pp. 166–172, 2018,[Online]. Available: http://ejournal. ust. ac. id/index. php/Jurnal_Means*.

* Corresponding author