

PERBANDINGAN ALGORITMA C5.0 DAN REGRESI LINEAR UNTUK PREDIKSI KELULUSAN MAHASISWA

Petti Sijabat¹⁾, Agustina Simangunsong²⁾

^{1,2} Program Studi Teknik Informatika STMIK Pelita Nusantara, Medan, Indonesia
email: sijabatpetti@yahoo.com, simangunsongagus@gmail.com

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 93.72% (testing) while the R2 score from the Linear Regression reached 33.31% (training) and 40.30% (testing).

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

1. PENDAHULUAN

Dengan semakin berkembangnya kemajuan teknologi dan didukung oleh pengetahuan manusia memberi pengaruh yang sangat baik terhadap teknologi penyimpangan data dan informasi. Dengan berkembangnya teknologi disegala bidang sehingga memperoleh banyak informasi dalam bentuk data berupa bidang industri ekonomi, dan lain-lain. Informasi-informasi yang diperoleh sangat banyak akan tetapi data-data tersebut masih banyak yang belum dimanfaatkan dengan baik.

Data tersebut diolah dengan menganalisis data yang dinamakan data mining, konsep data ini merupakan proses menemukan pola data terpilih, dengan menggunakan metode. *Knowledge Discovery in Database* (KDD) merupakan proses ringkasan informasi yang penting dari basis data yang besar. Data mining membagi model data yang diperoleh pada kumpulan data tersebut, sehingga data tersebut dapat lebih dioptimalkan penggunaanya di kehidupan sehari-hari [1]. Big data yang merupakan data berukuran besar yang terkumpul berupa set data yang sangat berharga dan dapat digunakan proses analisis, dimana hasil yang

diperoleh dapat berbentuk pengetahuan maupun informasi sehingga dapat dipergunakan untuk saat ini sebagai acuan perbandingan dengan data sebelumnya. Proses ini tidak hanya berlaku didunia bisnis tetapi dapat diterapkan pada instansi seperti perguruan tinggi yang memiliki banyak data [2].

Pada perguruan tinggi, terdapat banyak data mahasiswa dan data jumlah kelulusan disetiap tahunnya yang mampu menghasilkan informasi. salah satu parameter prestasi dari suatu perguruan tinggi adalah presentasi kelulusan mahasiswa yang lulus tepat sesuai dengan peraturan badan akreditasi nasional perguruan tinggi nomor 3 tahun 2019.

Salah satu perguruan tinggi yang memiliki banyak data yaitu Universitas Prima Indonesia dimana salah satunya yaitu fakultas teknologi dan ilmu komputer. Fakultas teknologi dan ilmu komputer merangkum data mahasiswa di dalam sistem informasi akademik mahasiswa yang memiliki database dan mengelolanya sesuai dengan apa dibutuhkan. Berdasarkan data yang diperoleh, Terdapat bahwa fakultas teknologi dan ilmu komputer, memiliki jumlah mahasiswa sebanyak 174 di tahun 2017, dan yang

lulus tepat waktunya sebanyak 156 orang, sehingga didapat persentasi kelulusan sebesar 89%. Pada tahun 2018 sebanyak 268 orang, diperoleh lulus tepat waktu yaitu 240 mahasiswa, dan didapatkan persentasi kelulusan 89%, sedangkan ditahun 2020 terdapat 273 mahasiswa, dan yang lulus tepat waktu 254 mahasiswa, sehingga didapat persentasi kelulusan mahasiswa 93%.

Sesuai dengan hasil dari pendataan peneliti dan hasil data-data yang ada bahwa setiap penerimaan mahasiswa baru jumlah yang diterima semakin bertambah, tetapi tidak semua mahasiswa dapat lulus tepat waktu yang akan mengakibatkan penumpukkan data sehingga data yang masuk dengan data yang keluar tidak sinkron. Sehingga mengakibatkan peningkatan jumlah data dan data akademis seluruh fakultas dari semua mahasiswa yang masih terdaftar setiap tahunnya.

Ada juga penelitian telah dilaksanakan untuk memprediksi kelulusan mahasiswa tepat waktu di universitas lainnya dengan penerapan algoritma- algoritma yang sudah ada. Salah satunya adalah *decision tree*. Terdapat beberapa kasus pula yang telah dilakukan seperti prediksi kelulusan tepat waktu mahasiswa menggunakan algoritma naive bayes [3]. Ada juga enerapan teknik klasifikasi untuk prediksi kelulusan mahasiswa dengan penerapan algoritma K-Nearest Neighbour [4], ada juga peneliti membuat perbandingan algoritma C4.5 dan naive bayes dengan kasus lulusan tepat waktu mahasiswa [5]. Analisa memprediksi dalam kasus mahasiswa drop out dengan penerapan metode decision tree dengan Algoritma ID3 dan C4.5 [6], ada juga kasus dalam prediksi kelulusan mahasiswa menggunakan metode decision tree dengan menggunakan algoritma C.45 [7]. Adanya hasil-hasil penelitian yang sudah ada dan adanya masalah-masalah dalam ruang lingkup fakultas tentang jumlah mahasiswa maka tim peneliti tertarik untuk melakukan suatu penelitian yaitu prediksi kelulusan mahasiswa menggunakan perbandingan algoritma C5.0 dengan *regression linear* yang akan digunakan dalam memprediksi kelulusan mahasiswa yang akan diuji coba pada mahasiswa fakultas teknologi informasi dan ilmu komputer. Sesuai dengan hasil penelitian diharapkan dapat menjadi

tolak ukur untuk meningkatkan kualitas mahasiswa di Universita Prima Indonesia.

2. METODE PENELITIAN



Gambar 2.1 Alur Metodologi Penelitian

Tahap awal peneliti mengumpulkan data-data yang dibutuhkan untuk proses data mining. Dengan menggunakan dari fakultas teknologi dan ilmu komputer. Data Mahasiswa/Alumni yang diambil memiliki jumlah 715 baris serta 9 atribut/kolom dari tahun 2017-2020. Data hasil seleksi ini akan digunakan dalam memprosesan data.

Tahap awal yang dilakukan sebelum menganalisis yaitu tahap data processing, perlu dilakukan suatu analisis terhadap data yang sudah diambil untuk dapat memeriksa kejelasan dan kelengkapan data penelitian. Tujuan dari analisis data yaitu untuk mengetahui kualitas data yang akan diuji serta dapat menarik hipotesis terkait dengan data yang akan diuji. Adapun beberapa langkah dalam menganalisis data yaitu: (a) data Preprocessing, pada tahap ini adalah tahap untuk pembersihan data (*data cleaning*). Pembersihan data yang dilakukan dalam penelitian ini adalah menghapus atribut-atribut data yang tidak dibutuhkan dalam proses prediksi nantinya serta menghapus data-data yang bernilai null atau N/A. (b) data transformation, tahap berikutnya adalah tahap data transformation. Pada tahap ini akan dilakukan proses normalisasi data. Tujuan dari normalisasi ini agar data yang khususnya bersifat numerik berada pada range 0 sampai 1 sehingga persebaran data tidak terlalu jauh. Sedangkan untuk data kategorikal akan diubah menjadi

numerik sehingga dapat dicocokkan dalam proses data mining selanjutnya. (c) data mining, merupakan proses mencari pola telah dipilih dan ditransformasi sesuai dengan algoritma C5.0 dan regression linear, yang nanti algoritma ini memiliki nilai perbandingan terhadap prediksi kelulusan tepat waktu mahasiswa teknologi informasi dan ilmu komputer pada tahun 2017-2020 berdasarkan total jumlah mahasiswa yang terdata.

Uji Coba Python merupakan tahap implementasi analisis dan prediksi Kelulusan Mahasiswa Fakultas Teknologi Informasi dan Ilmu Komputer di Universitas Prima dimana akan dilakukan pengujian berupa program yang ditulis dengan Python versi 3.7.12 dalam platform Google Colab. Python disebut sebagai bahasa pemrograman yang mampu menggabungkan keterampilan serta kemampuan dengan menggunakan tata bahasa melalui kode yang jelas, untuk memberi perintah pada program yang dibuat[8].

Pengujian sistem merupakan tahap pengekskusi sistem perangkat lunak untuk menentukan apakah penerapan Data Mining ini cocok dengan spesifikasi sistem. Pengujian sistem ini berupa memverifikasi apakah program yang telah dibuat dapat dijalankan sesuai dengan spesifikasi dan rancangan yang sudah dibuat.

Tahap Evaluasi ini merupakan tahap terakhir yang dilakukan untuk menghasilkan kapasitas dan efektivitas model yang mampu memenuhi tujuan dan dapat memberi solusi pada saat memecahkan masalah dipenelitian ini, serta mengambil keputusan berkaitan dengan penggunaan hasil dari Data Mining.

3. HASIL DAN PEMBAHASAN

3.1 Analisis Data

3.1.1 Dataset Kelulusan Mahasiswa

Peneliti menggunakan dua metode yaitu algoritma C5.0 dan regression linear terhadap prediksi kelulusan mahasiswa fakultas teknologi dan ilmu komputer yang diimplementasikan ke dalam sebuah program berbasis Phyton. Adapun data yang dipakai merupakan data kelulusan

mahasiswa pada tahun 2017- 2020, dapat dilihat pada Tabel 3.1.

Tabel 3.1 Data sampel kelulusan mahasiswa

NIM	Nama	Pro di	No. Ijazah	IPK	JK	Lama Studi
12330 30300 78	Suba nto	TI	294/S1- TI/FTIK/ UNPRI/I V/2017	3.18	LK	Terla mbat
12330 30300 84	Andr e	TI	296/S1- TI/FTIK/ UNPRI/I V/2017	2.67	LK	Terla mbat
12330 30300 08	Nove ndy	TI	297/S1- TI/FTIK/ UNPRI/I V/2017	3.49	LK	Tepat
15330 30402 41	Desi Gulo	SI	654/S1- SI/FTIK/ UNPRI/ VI/2020	2.47	PR	Terla mbat
16330 30304 08	Cindy	SI	779/S1- SI/FTIK/ UNPRI/I X/2020	3.70	PR	Tepat
16330 30104 06	Vane ssa	TIN D	68/S1- TIND/FT IK/UNP RI/IX/20 20	3.82	PR	Tepat
16330 30104 07	Wilso n	TIN D	69/S1- TIND/FT IK/UNP RI/IX/20 20	3.68	LK	Tepat
16330 30204 05	Wand y	TE	39/S1- TE/FTIK /UNPRI/ IX/2020	3.70	LK	Tepat
15330 30304 44	Isnain i	SI	771/S1- SI/FTIK/ UNPRI/I X/2020	3.34	PR	Terla mbat
16330 30304 31	Steve n	SI	749/S1- SI/FTIK/ UNPRI/I X/2020	3.51	LK	Tepat

Pengambilan data kelulusan ini dimulai dari tahun kelulusan 2017 sampai 2020. Berdasarkan data yang diterima oleh peneliti dari fakultas teknologi dan ilmu komputer, data tersebut dapat terkumpul menjadi sebuah data set yang diteliti. Didalam dataset ini ada beberapa atribut yang dipakai yaitu NIM, Nama, Prodi, No. Alumni, Nomor Ijazah, IPK, Jenis Kelamin, Lama Studi.

3.1.2 Analisis Kelulusan Mahasiswa

Data yang telah diperoleh akan dianalisis dan dibuat visualisasi terhadap beberapa variabel sehingga bisa dijadikan sebagai informasi yang berguna.

3.1.3 Analisis Jumlah Mahasiswa berdasarkan Tahun Ajaran

Disini Analisis yang dilakukan terhadap jumlah mahasiswa berdasarkan disetiap program studi (prodi) berdasarkan tahun ajaran, sesuai gambar dibawah jumlah mahasiswa pada prodi Teknik Informatika (61), Sistem Informasi (91), Teknik Industri (17), dan Teknik Elektro (5) pada Tahun 2017. Pada tahun 2018 Prodi Teknik Informatika (134), Sistem Informasi (105), Teknik Industri (19), dan Teknik Elektro (10), sedangkan pada Tahun 2020 Prodi Teknik Informatika (109), Sistem Informasi (118), Teknik Industri (30), dan Teknik Elektro (15).



Gambar 3.1 Analisis Jumlah Mahasiswa Berdasarkan Tahun Ajaran

3.1.4 Analisis Jumlah Mahasiswa berdasarkan Jenis Kelamin

Pada tahap ini analisis yang dilakukan berdasarkan Jenis Kelamin mahasiswa yang terdapat di Fakultas Teknologi dan Informasi Universitas Prima Indonesia. Pada Tabel 3.2. jumlah mahasiswa berdasarkan jenis kelamin

laki-laki berjumlah 480 mahasiswa sedangkan perempuan berjumlah 235 mahasiswa.

Tabel 3.2 Jumlah Mahasiswa berdasarkan Jenis Kelamin

Jenis Kelamin (JK)	Jumlah Mahasiswa
Laki – laki (L)	480
Perempuan (P)	235

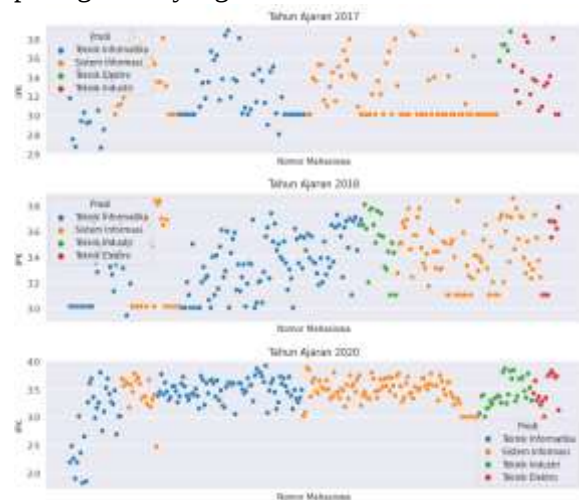
Sedangkan pada gambar dibawah ini merupakan visualisasi atau tampilan dari jumlah mahasiswa berdasarkan Jenis Kelamin dengan tampilan berbentuk Persen (%), yang dimana Jenis Kelamin Laki-laki 67% sedangkan Perempuan 33%.



Gambar 3.2 Visualisasi Jumlah Mahasiswa berdasarkan Jenis Kelamin

3.1.5 Analisis IPK Mahasiswa Berdasarkan Program Studi

Pada tahap ini dilakukan analisis terhadap Nilai IPK (Indeks Prestasi Kumulatif) Mahasiswa yang dimana berdasarkan Program Studi mahasiswa dari tahun 2017-2020. Seperti pada gambar yang tertera dibawah.

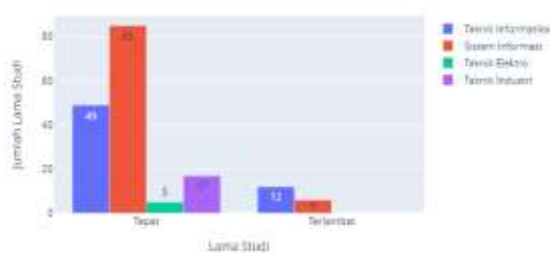


Gambar 3.3 Analisis IPK Mahasiswa berdasarkan Program Studi

3.1.6 Analisis Kelulusan Mahasiswa Berdasarkan Lama Studi

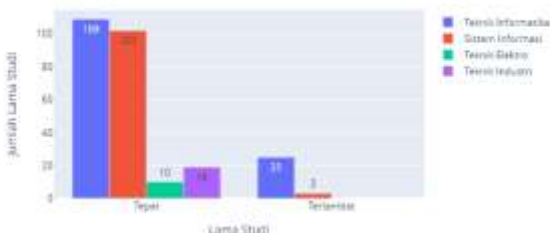
Pada tahap Analisis ini akan di tampilkan visualisasi kelulusan jumlah mahasiswa berdasarkan lama studi terhadap program studi masing-masing. Seperti pada gambar dibawah dapat dilihat jumlah mahasiswa yang tepat lulus dantidak tepat lulus (terlambat) pada tahun 2017, 2018 dan 2020.

Jumlah Lama Studi Tahun Ajaran 2017



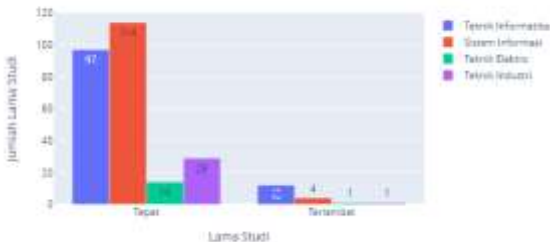
Gambar 3.4 Analisis Kelulusan Mahasiswa Berdasarkan Lama Studi Tahun 2017

Jumlah Lama Studi Tahun Ajaran 2018



Gambar 3.5 Analisis Kelulusan Mahasiswa Berdasarkan Lama Studi Tahun 2018

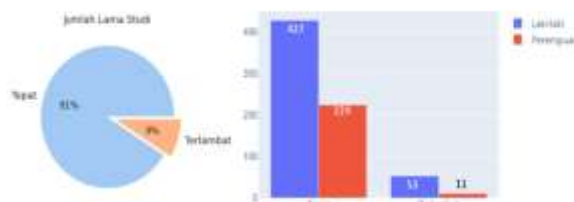
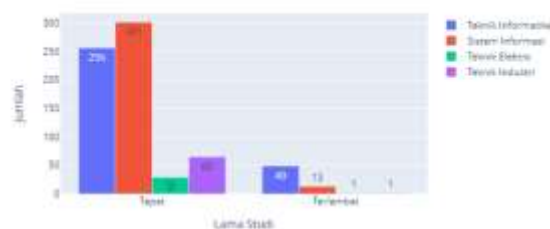
Jumlah Lama Studi Tahun Ajaran 2020



Gambar 3.6 Analisis Kelulusan Mahasiswa Berdasarkan Lama Studi Tahun 2020

Dari hasil visualisasi diatas dapat disimpulkan jumlah keseluruhan kelulusan mahasiswa yang tepat tulus berjumlah 651 (91%) dengan jenis kelamin laki-laki berjumlah 427 mahasiswa dan jenis kelamin perempuan berjumlah 224 mahasiswa sedangkan yang tidak tepat lulus (terlambat) berjumlah 64 (9%) dengan jenis kelamin laki-laki berjumlah 53 mahasiswa dan jenis kelamin perempuan berjumlah 11 mahasiswa. Seperti pada gambar yang akan tampil.

Grafik Mahasiswa Berdasarkan Lama Studi



Gambar 3.7 Analisis Jumlah Keseluruhan Kelulusan Mahasiswa Berdasarkan Lama Studi Tahun 2017-2020

3.1.7 Transformasi data pada dataset

Dimana dikatakan bahwa pengertian dari transformasi data adalah proses penggantian data ke dalam bentuk yang sesuai untuk diproses pada data mining, dikarenakan terkadang data yang digunakan dalam pemrosesan data mining mempunyai bentuk yang belum sesuai sehingga tidak dapat langsung digunakan. Dengan sistem tersebut perlu dirubah bentuknya. Hasil dari perubahan format yang dibentuk dalam tabel yang telah dikelompokkan sesuai dengan target yang akan diprediksi yaitu kelulusan mahasiswa, dalam tahap transformasi data membutuhkan beberapa variabel- variabel, antara lain yaitu variabel prediktor: NIM, Nama, Prodi, No. Ijazah, IPK, Jenis kelamin, Lama Studi, Tahun

Ajaran. Variabel Kriteria: “Tepat” atau “Terlambat”.

Tabel 3.3 Transformasi Dataset

NIM	Nama Prodi	No. Ijazah	IPK	JK	Lama Studi	Tahun Ajaran	
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 Uji Coba dengan Menggunakan Algoritma C5.0

Setiap algoritma yang sudah diuji memiliki nilai yang signifikan dan memiliki ciri dari algoritma masing-masing. Algoritma C5.0 merupakan penyempurnaan algoritma dari ID3 dan C4.5. Dimana algoritma C5.0 dikatakan mampu memberikan akurasi data yang sangat baik, dimana tingkat kesalahan yang rendah pada kasus yang tidak terlihat. Dan juga secara otomatis dapat menghapus data yang tidak penting. Dalam algoritma ini pemilihan atribut diproses menggunakan *gain ration* [9]. Jika, perhitungan C4.5 hanya sampai menghitung gain, sedangkan pada algoritma C5.0 akan melanjutkan perhitungan *gain ration* dengan *gain* dan *entropy* yang sudah ditentukan.

Adapun rumus untuk mencari nilai Entropy dan Gain adalah sebagai berikut :

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

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

Dimana Entropy: S: Himpunan kasus, n: umlah partisi S, pi : Proporsi dari Si dan S. dan Gain : S : Himpunan kasus, Si : Himpunan kasus pada kategori ke-i, A : Variabel, m : Jumlah

kategori pada variabel A, Si: Jumlah kasus pada kategori ke-I, S:Jumlah kasus dalam S

Dari percobaan diatas didapatkan hasil uji coba Prediksi Kelulusan Mahasiswa dengan Algoritma C5.0 sebagai Berikut

Tabel 3.4 Uji Coba Algoritma C5.0

Actual	Predicted
0	1
0	0
0	0
0	1
1	1
0	0
0	0
0	0
1	1
0	0

3.3 Uji Coba dengan Menggunakan Regression Linear

Regression Linear adalah Algoritma yang melibatkan hubungan antara satu variabel tak bebas digabungkan dnegan satu variabel bebas. Regression Linear berfungsi untuk mengetes sejauh mana hubungan sebab- akibat antara variabel faktor (X) terhadap variabel akibatnya (Y). Regression linear dengan satu variabel prediktor disebut Regression Sederhana, sedangkan untuk variabel prediktor lebih dari satu disebut Regression Linear [10].

Pada penelitian ini, algoritma yang digunakan yaitu Regresion Linear. Untuk cara kerja Reggresion Linear Berganda sendiri dengan meramalkan pengaruh dari kedua variabel prediktor atau lebih X1,...,Xn terhadap satu variabel Y agar dapat membuktikan kebenaran adanya hubungan antara dua variabel input atau lebih X1,...,Xn dengan satu variabel Y [1].

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

Dimana Regression Linear: y :Variabel Kriteria, w_0 : Bias, w_(i) : Koefisien, x_(i) :Variabel Prediktor, n :Jumlah Data.

Dari uji coba diatas didapatkan hasil Prediksi Kelulusan Mahasiswa dengan

menggunakan Algoritma Regresion Linear. Berikut hasil uji coba dengan menggunakan Algoritma Regression Linear.

Tabel 3.5 Uji Coba Algoritma Regression Linear

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

3.4 Evaluasi Nilai Algoritma C5.0 dengan Regression Linear

Tahap selanjutnya adalah Tahap Evaluasi. Tahap ini merupakan tahap dimana hasil- hasil dari uji coba tersebut di evaluasi dengan menggunakan nilai R2 Score. Nilai pada R2 digunakan sebagai kuantitas untuk memperkirakan persentasi variance dari masing-masing algoritma.

$$\text{Accuracy} = \frac{TP + FP}{TP + FP + TN + FN}$$

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

Dari hasil Evaluasi diatas, dapat dibuatkan hasil Perbandingan R2 Score antara Algoritma C5.0 dengan Regression Linear.

Tabel 3.6 Hasil Perbandingan Akurasi Training dan Testing pada Algoritma C5.0 dengan Regression Linear

Model	Akurasi Training (%)	Akurasi Testing (%)
Algoritma C5.0	96.85	93.72
Regression Linear	33.31	40.30

4. KESIMPULAN

Dari hasil penelitian dan uji coba data, peneliti menuliskan ada beberapa kesimpulan pada penelitian ini antara lain:

1. Berdasarkan lama studi, mahasiswa yang tepat lulus berjumlah 651 (91%) dengan jenis kelamin laki-laki berjumlah 427 mahasiswa dan jenis kelamin perempuan (P) berjumlah 224 mahasiswa sedangkan yang tidak tepat lulus (terlambat) berjumlah 64 (9%) dengan jenis kelamin laki-laki (L) berjumlah 53 mahasiswa dan jenis kelamin perempuan berjumlah 11 mahasiswa dari tahun 2017-2020.
2. Hasil perbandingan R2 score dari algoritma C5.0 mencapai 96.85% (training) dan 93.72% (testing) sedangkan hasil R2 score dari Regression Linear mencapai 33.31% (training) dan 40.30% (testing).

5. REFERENSI

- [1] D. S. O. Panggabean, E. Buulolo, and N. Silalahi, "Penerapan Data Mining Untuk Memprediksi Pemesanan Bibit Pohon Dengan Regresi Linear Berganda," *JURIKOM (Jurnal Ris. Komputer)*, vol. 7, no. 1, p. 56, 2020, doi: 10.30865/jurikom.v7i1.1947.
- [2] E. Setiawan, D. Antoni, and A. H. Mirza, "Analisis Penerimaan Sistem Ujian Online Berbayar Dengan Menggunakan Metode Technology Acceptance Model (Tam) Dan Webqual," *J. Bina Komput.*, vol. 1, no. 1, pp. 61–72, 2019, doi: 10.33557/binakomputer.v1i1.155.
- [3] A. A. Murtopo, "Prediksi Kelulusan Tepat Waktu Mahasiswa STMIK YMI Tegal Menggunakan Algoritma Naïve Bayes," *CSRID (Computer Sci. Res. Its Dev. Journal)*, vol. 7, no. 3, p. 145, 2016, doi: 10.22303/csrid.7.3.2015.145-154.
- [4] A. Y. Saputra and Y. Primadasa, "Penerapan Teknik Klasifikasi Untuk Prediksi Kelulusan Mahasiswa Menggunakan Algoritma K-Nearest Neighbor," *Techno.Com*, vol. 17, no. 4,



- pp. 395–403, 2018, doi: 10.33633/tc.v17i4.1864.
- [5] M. Kamil and W. Cholil, “Analisis Perbandingan Algoritma C4.5 dan Naive Bayes pada Lulusan Tepat Waktu Mahasiswa di Universitas Islam Negeri Raden Fatah Palembang,” *J. Inform.*, vol. 7, no. 2, pp. 97–106, 2020, doi: 10.31294/ji.v7i2.7723.
- [6] L. R. Haidar, E. Sedyono, and A. Iriani, “Analisa Prediksi Mahasiswa Drop Out Menggunakan Metode Decision Tree Dengan Algoritma ID3 dan C4.5,” *J. Transform.*, vol. 17, no. 2, p. 97, 2020, doi: 10.26623/transformatika.v17i2.1609.
- [7] Mashlahah, *Prediksi Kelulusan Mahasiswa Menggunakan Metode Decision Tree Dengan Penerapan Algoritma C4.5*, 2013th ed. Malang: Universitas Islam Negeri Maulana Malik Ibrahim, 2013.
- [8] M. M. Baharuddin, H. Azis, and T. Hasanuddin, “Analisis Performa Metode K-Nearest Neighbor Untuk Identifikasi Jenis Kaca,” *Ilk. J. Ilm.*, vol. 11, no. 3, pp. 269–274, 2019, doi: 10.33096/ilkom.v11i3.489.269-274.
- [9] K. P. Wirdhaningsih, M. Ratnawati, Dian Eka, U. B. Malang, D. Mining, and D. Tree, *Penerapan Algoritma Decision Tree C5.0 Untuk Peramalan Forex*, 2013th ed. Malang: Universitas Brawijaya Malang, 2013.
- [10] H. W. Herwanto, T. Widiyaningtyas, and P. Indriana, “Penerapan Algoritme Linear Regression untuk Prediksi Hasil Panen Tanaman Padi,” *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 8, no. 4, p. 364, 2019, doi: 10.22146/jnteti.v8i4.537.