

BAB IV

HASIL DAN PEMBAHASAN

4.1. Hasil

Data yang digunakan dalam proses klasifikasi diperoleh dari proses pengumpulan data dokumentasi historis akademik Biro Administrasi Akademik dan Kemahasiswaan (BAAK) Universitas Muhammadiyah Pringsewu. Data yang akan diolah dalam pembahasan ini merupakan data mahasiswa angkatan 2016 dan 2017 program sarjana. Tujuan dari penelitian ini adalah untuk membandingkan model algoritma yang memiliki akurasi sangat tinggi menggunakan metode algoritma *Decision Tree C4.5* dan *Naive Bayes*. Data yang digunakan berjumlah 881 terdiri dari 9 atribut dengan 1 atribut prediksi. Untuk atribut yang digunakan dapat dilihat pada tabel 4.1 dibawah ini :

Tabel 4.1 Atribut Data yang digunakan

No	Atribut	Tipe	Keterangan
1	NIM	Integer	16010001
2	Program Studi	Polynominal	Program Studi
3	Jenis Kelamin	Polynominal	L/P
4	Status Pernikahan	Polynominal	Menikah / Belum Menikah
5	Status Pekerjaan	Polynominal	Sudah Bekerja / Belum Bekerja
6	IPK Semester 5	Real	2,00 – 4,00
7	Asal Mahasiswa	Polynominal	Pringsewu / Luar Pringsewu
8	SKS Total	Polynominal	80 sks – 112 sks
9	Keterangan	Binominal	Tidak Tepat Waktu / Tepat Waktu

4.2 Pembahasan

4.2.1 Persiapan Data

Pengolahan data awal yang dilakukan pada *data mining* untuk mengubah data mentah (*raw data*) yang dikumpulkan dari berbagai sumber menjadi informasi yang lebih bersih yang kemudian digunakan untuk proses pengolahan data selanjutnya. Berdasarkan sebaran data diperoleh hasil *raw data* sebanyak 881 data. Potongan *dataset* bisa dilihat pada gambar 4.1 berikut:

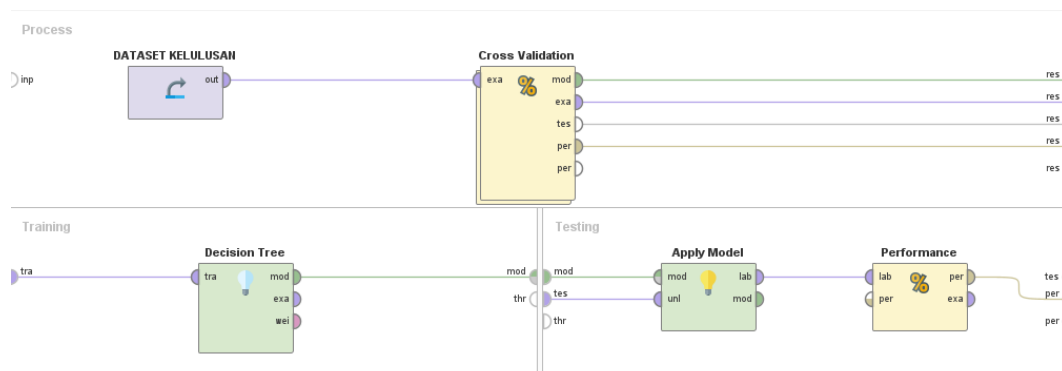
Row No.	NIM	Keterangan	Program St...	L/P	Status Perni...	Status Peke...	IPK Sem 5	SKS Total	Asal Mahasi...
1	16010001	Tidak Tepat ...	S1 Manajemen	L	Belum	Tidak	3.220	104	Luar Pringse...
2	16010002	Tepat Waktu	S1 Manajemen	P	Belum	Tidak	3.220	104	Pringsewu
3	16010003	Tepat Waktu	S1 Manajemen	P	Belum	Tidak	3.530	107	Pringsewu
4	16010004	Tepat Waktu	S1 Manajemen	P	Belum	Tidak	3.350	107	Pringsewu
5	16010005	Tepat Waktu	S1 Manajemen	L	Belum	Tidak	3.150	104	Pringsewu
6	16010006	Tepat Waktu	S1 Manajemen	L	Belum	Tidak	3.250	104	Pringsewu
7	16010007	Tepat Waktu	S1 Manajemen	P	Belum	Tidak	3.700	107	Pringsewu
8	16010008	Tepat Waktu	S1 Manajemen	P	Belum	Tidak	3.480	107	Luar Pringse...
9	16010009	Tepat Waktu	S1 Manajemen	P	Belum	Tidak	3.630	107	Pringsewu
10	16010010	Tepat Waktu	S1 Manajemen	P	Belum	Tidak	3.520	107	Pringsewu
11	16010011	Tepat Waktu	S1 Manajemen	P	Belum	Tidak	3.370	107	Luar Pringse...
12	16010012	Tepat Waktu	S1 Manajemen	L	Belum	Tidak	3.510	107	Luar Pringse...
13	16010013	Tepat Waktu	S1 Manajemen	P	Belum	Tidak	3.640	107	Pringsewu
14	16010014	Tepat Waktu	S1 Manajemen	L	Belum	Tidak	3.210	107	Pringsewu
15	16010015	Tepat Waktu	S1 Manajemen	L	Belum	Tidak	3.150	104	Pringsewu

Gambar 4.1 Potongan Dataset

4.2.2 Pemodelan

1. Proses Klasifikasi Menggunakan Algoritma Decision Tree C4.5

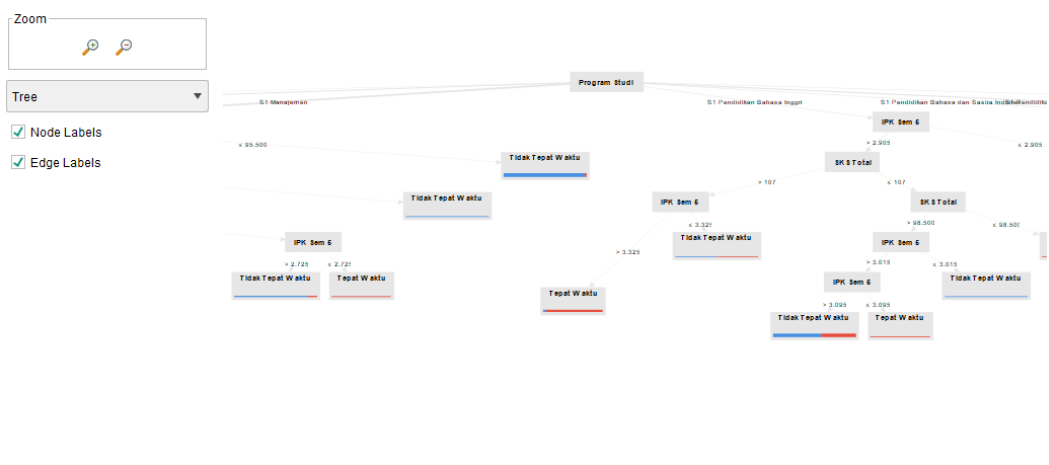
Proses ini merupakan implementasi pembuatan model klasifikasi pada pengklasifikasian data. Pada proses ini terdapat dua tahap yaitu pembentukan *tree* dan perubahan *tree* menjadi *rule*. Pada proses ini digunakan aplikasi Rapid Miners sebagai alat bantu untuk membuat proses *data mining*. Pada algoritma *Decision Tree*, *record* yang sudah diimpor ke Rapid Miner digunakan untuk menentukan pola pohon keputusan. Penerapan data pada Rapid Miner digunakan untuk Prediksi kelulusan mahasiswa menggunakan algoritma *Decision Tree* ditunjukkan pada gambar 4.2 dibawah ini:



Gambar 4.2 Model Klasifikasi Algoritma C4.5

Gambar 4.2 menunjukkan proses penerapan klasifikasi menggunakan *tool* Rapid Miner. Data yang sudah disiapkan diimpor ke Rapid Miner untuk melakukan pengujian menggunakan *Cross Validation* untuk membagi data menjadi data *training* dan *data testing* karena data yang digunakan adalah *supervised* dan algoritma yang digunakan adalah algoritma *Decision Tree*. Tahapan tersebut merupakan tahapan pengolahan data tahap akhir dalam proses klasifikasi menggunakan model Algoritma C4.5, pada tahap penerapan model data yang diolah dapat menjadi informasi berupa model yang terbentuk dari proses pengklasifikasian.

Setelah melakukan beberapa langkah diatas dalam proses klasifikasi metode algoritma C4.5 maka akan diperoleh model yang terbentuk dari proses pengklasifikasian algoritma C4.5 berupa pohon keputusan seperti gambar 4.3 dibawah ini.



Gambar 4.3 Pohon Keputusan

Gambar 4.3 merupakan gambar pohon keputusan yang merupakan *output* dari proses klasifikasi menggunakan algoritma C4.5. pohon keputusan terbentuk berdasarkan *node*. *Node* dalam pohon keputusan merupakan variabel-variabel yang digunakan dalam penelitian. Berdasarkan gambar 4.3 maka faktor dari variabel yang paling dominan dalam mempengaruhi klasifikasi kelulusan mahasiswa yaitu Program Studi. Selain pohon keputusan seperti pada gambar 4.3 juga terdapat *text view* yang menjelaskan berdasarkan pohon keputusan seperti pada gambar 4.4 dibawah ini.

Tree

```

Program Studi = S1 Bimbingan dan Konseling
| SKS Total > 109
| | SKS Total > 112
| | | IPK Sem 5 > 3.275
| | | | IPK Sem 5 > 3.480: Tidak Tepat Waktu {Tidak Tepat Waktu=2, Tepat Waktu=2}
| | | | IPK Sem 5 ≤ 3.480: Tepat Waktu {Tidak Tepat Waktu=0, Tepat Waktu=8}
| | | | IPK Sem 5 ≤ 3.275: Tidak Tepat Waktu {Tidak Tepat Waktu=2, Tepat Waktu=1}
| | | SKS Total ≤ 112: Tepat Waktu {Tidak Tepat Waktu=0, Tepat Waktu=25}
| | SKS Total ≤ 109: Tidak Tepat Waktu {Tidak Tepat Waktu=7, Tepat Waktu=1}
Program Studi = S1 Keperawatan: Tepat Waktu {Tidak Tepat Waktu=8, Tepat Waktu=145}
Program Studi = S1 Manajemen
| SKS Total > 95.500
| | IPK Sem 5 > 2.505
| | | IPK Sem 5 > 2.930
| | | | Jenis Kelamin = L
| | | | | IPK Sem 5 > 2.965
| | | | | IPK Sem 5 > 3.020: Tepat Waktu {Tidak Tepat Waktu=16, Tepat Waktu=67}
| | | | | IPK Sem 5 ≤ 3.020: Tidak Tepat Waktu {Tidak Tepat Waktu=2, Tepat Waktu=0}
| | | | | IPK Sem 5 ≤ 2.965: Tepat Waktu {Tidak Tepat Waktu=0, Tepat Waktu=4}
| | | | | Jenis Kelamin = P
| | | | | Asal Mahasiswa = Luar Pringsewu
| | | | | IPK Sem 5 > 3.470
| | | | | IPK Sem 5 > 3.690
| | | | | IPK Sem 5 > 3.715: Tepat Waktu {Tidak Tepat Waktu=0, Tepat Waktu=3}

```

Gambar 4.4 Text View pohon keputusan

Gambar 4.4 merupakan *text view* pohon keputusan merupakan aturan (*rule*) yang didapatkan berdasarkan perhitungan dan pengujian data pada setiap atribut menggunakan algoritma C4.5. Selain pohon keputusan seperti pada gambar 4.3 dan *Text View* pohon keputusan seperti gambar 4.4 pengujian data menggunakan algoritma C4.5 juga diperoleh tabel hasil akurasi seperti pada gambar 4.5 dibawah ini Hasil pengujian dapat kita lihat pada gambar 4.5 dibawah ini.

accuracy: 79.91% +/- 5.26% (micro average: 79.91%)

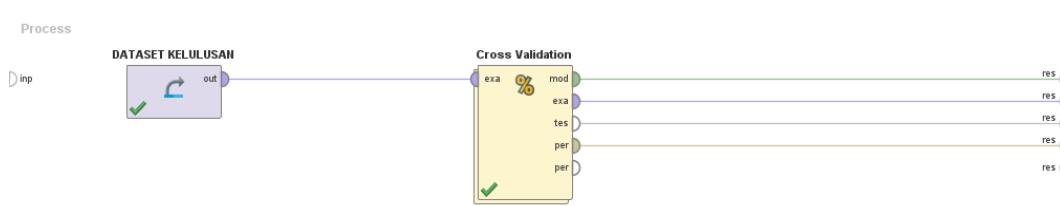
	true Tidak Tepat Waktu	true Tepat Waktu	class precision
pred. Tidak Tepat Waktu	201	115	63.61%
pred. Tepat Waktu	62	503	89.03%
class recall	76.43%	81.39%	

Gambar 4.5 Hasil Akurasi Pengujian Algoritma C4.5

Dari gambar 4.16 dapat disimpulkan bahwa tingkat akurasi metode algoritma C4.5 sangat tinggi yaitu mencapai 79,91%, dimana jumlah data yang diprediksi tidak tepat waktu dan kenyataannya tidak tepat waktu sebanyak 201, jumlah data yang diprediksi tepat waktu dan kenyataannya tidak tepat waktu sebanyak 115, jumlah data yang diprediksi tidak tepat waktu dan kenyataannya tepat waktu sebanyak 62, dan jumlah data yang diprediksi tepat waktu dan kenyataannya tepat waktu sebanyak 503.

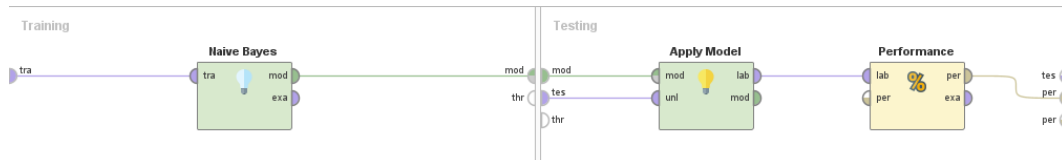
2. Proses Klasifikasi Menggunakan Algoritma *Naive Bayes*

Proses klasifikasi menggunakan model *Naive Bayes* digunakan untuk menggambarkan atau memprediksi peluang berdasarkan masing-masing kondisi. Pada proses ini digunakan aplikasi Rapid Miner sebagai alat bantu untuk membuat proses data mining. Berikut adalah gambaran penerapan model *Naive Bayes* menggunakan Rapid Miner.



Gambar 4.6 Skema Perubahan Target Label

Gambar 4.6 merupakan proses target label yang digunakan dalam penelitian ini yaitu nilai hasil evaluasi dengan label akhir “Lulus Tepat Waktu” dan “Lulus Tidak Tepat Waktu”. Setelah proses target label pada gambar 4.7 selanjutnya seperti pada gambar 4.7 di bawah ini.



Gambar 4.7 Model Klasifikasi Algoritma *Naive Bayes*

Gambar 4.7 merupakan proses pembentukan skema *Naive Bayes* dimana pada tahap tersebut dipilih suatu model *Naive Bayes* untuk melakukan pengujian terhadap data yang telah diinput. Berdasarkan gambar 4.7 yang telah dibangun pada aplikasi Rapid Miner maka diperoleh hasil pada gambar 4.8 sebagai berikut:

Attribute	Parameter	Tidak Tepat Waktu	Tepat Waktu
Program Studi	value=S1 Manajemen	0.289	0.322
Program Studi	value=S1 Bimbingan dan Konseling	0.042	0.060
Program Studi	value=S1 Pendidikan Matematika	0.106	0.071
Program Studi	value=S1 Pendidikan Bahasa dan Sastra...	0.023	0.123
Program Studi	value=S1 Pendidikan Bahasa Inggris	0.118	0.068
Program Studi	value=S1 Pendidikan Guru Sekolah Dasar	0.392	0.121
Program Studi	value=S1 Keperawatan	0.030	0.235
Program Studi	value=unknown	0.000	0.000
Jenis Kelamin	value=L	0.403	0.285
Jenis Kelamin	value=P	0.597	0.715
Jenis Kelamin	value=unknown	0.000	0.000
Status Pernikahan	value=Belum	0.996	1.000
Status Pernikahan	value=Sudah	0.004	0.000

Gambar 4.8 Hasil Pengujian *Naive Bayes*

Gambar 4.8 merupakan tabel distribusi hasil pengujian dari model *Naive Bayes*. Pada tabel tersebut diuraikan prediksi berdasarkan parameter dari masing-masing variabel. Selain tabel distribusi seperti pada gambar 4.8 hasil pengujian data menggunakan model *Naive Bayes* juga memperoleh hasil akurasi seperti pada gambar 4.9 dibawah ini.

accuracy: 76.95% +/- 2.37% (micro average: 76.96%)

	true Tidak Tepat Waktu	true Tepat Waktu	class precision
pred. Tidak Tepat Waktu	70	10	87.50%
pred. Tepat Waktu	193	608	75.91%
class recall	26.62%	98.38%	

Gambar 4.9 Hasil Akurasi Pengujian Algoritma *Naive Bayes*

Dari gambar 4.9 dapat dilihat bahwa pengujian data yang dilakukan dengan menggunakan model *Naive Bayes* memiliki tingkat akurasi yang cukup tinggi yaitu 76,95%, hal ini menunjukkan bahwa proses klasifikasi sudah baik. Dimana jumlah data yang diprediksi Tidak tepat waktu dan kenyataannya tidak tepat waktu sebanyak 70, jumlah data yang diprediksi Tepat waktu dan kenyataannya tidak tepat waktu sebanyak 10, jumlah data yang diprediksi tidak tepat waktu dan kenyataannya tepat waktu sebanyak 193, dan jumlah data yang diprediksi tepat waktu dan kenyataannya waktu sebanyak 608. Untuk gambaran *Simple Distribution* dari metode *Naive Bayes* dapat dilihat pada gambar 4.10 dibawah ini.

SimpleDistribution

Distribution model for label attribute Keterangan

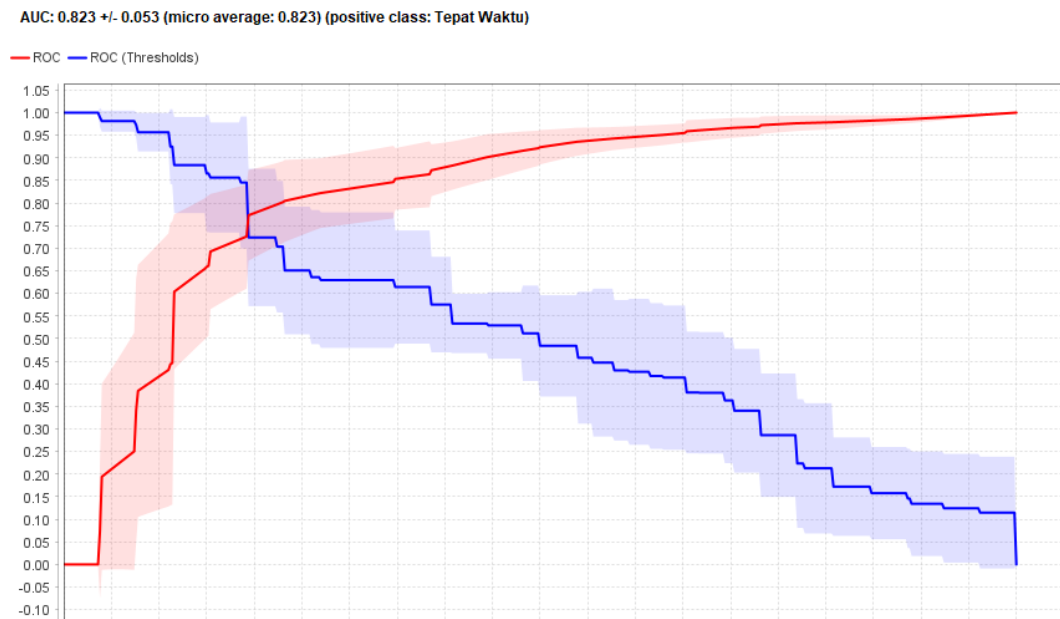
Class Tidak Tepat Waktu (0.299)
7 distributions

Class Tepat Waktu (0.701)
7 distributions

Gambar 4.10 *Simple Distribution Naive Bayes*

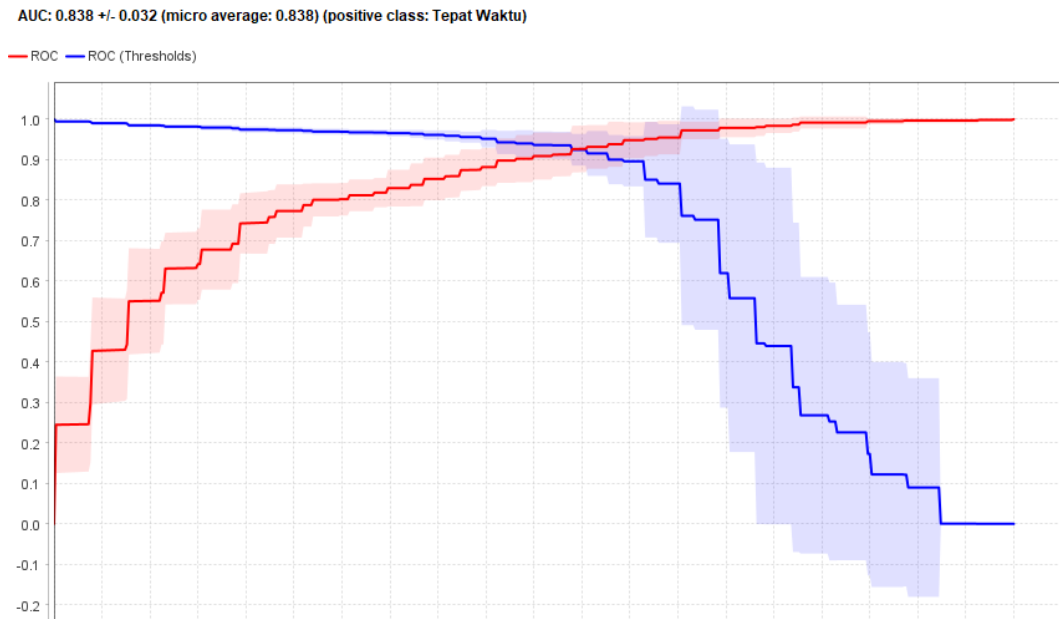
Hasil dari klasifikasi data kelulusan tepat waktu yang diuji menggunakan *Naive Bayes* dapat dilihat berdasarkan gambar 4.10 dimana dari data kelulusan tepat waktu terdapat dua label “Tidak Tepat Waktu atau Tepat Waktu” label “Tidak Tepat Waktu” menunjukkan kelas dengan nilai 0,299 dan label “Tepat Waktu” menunjukkan kelas dengan nilai 0,701.

Selain *Confusion Matrix* untuk mengetahui kinerja dari eksperimen ini kita dapat mengandalkan kurva AUC yang dihasilkan. Perbandingan hasil Kurva AUC menggunakan algoritma C4.5 dan *Naive Bayes* dapat kita lihat pada gambar 4.11 dan 4.12 dibawah ini:



Gambar 4.11 Kurva AUC C4.5

Kurva ROC pada gambar 4.11 menunjukkan hasil akurasi dan membandingkan klasifikasi secara visual dengan *false positif* sebagai garis horizontal dan *true positif* sebagai garis vertikal. Dari gambar 4.22 merupakan visualisasi dari hasil AUC 0,823 yang didapatkan Algoritma C4.5. Sedangkan untuk hasil ROC *Naive Bayes* dapat dilihat pada gambar 4.12 dibawah ini



Gambar 4.12 Kurva AUC Naive Bayes

Kurva ROC *Naive Bayes* yang ditunjukkan pada gambar 4.12 menunjukkan visualisasi dari hasil AUC 0.838 yang termasuk dalam kategori *Good Classification*.

4.3 Evaluasi

Berdasarkan pengujian yang telah dilakukan menggunakan 2 (dua) algoritma yaitu *Naive Bayes* dan Algoritma C4.5 dan menggunakan 881 data sebagai sampel pengujian untuk prediksi kelulusan tepat waktu mahasiswa Universitas Muhammadiyah Pringsewu didapatkan hasil sebuah model sebagai berikut:

1. Program Studi = S1 Bimbingan dan Konseling
 - | SKS Total > 109
 - | | SKS Total > 112
 - | | | IPK Sem 5 > 3.275
 - | | | | IPK Sem 5 > 3.480: Tidak Tepat Waktu {Tidak Tepat Waktu=2, Tepat Waktu=2}
 - | | | | IPK Sem 5 ≤ 3.480: Tepat Waktu {Tidak Tepat Waktu=0, Tepat Waktu=8}
 - | | | IPK Sem 5 ≤ 3.275: Tidak Tepat Waktu {Tidak Tepat Waktu=2, Tepat Waktu=1}
 - | | SKS Total ≤ 112: Tepat Waktu {Tidak Tepat Waktu=0, Tepat Waktu=25}
 - | SKS Total ≤ 109: Tidak Tepat Waktu {Tidak Tepat Waktu=7, Tepat Waktu=1}
2. Program Studi = S1 Keperawatan: Tepat Waktu {Tidak Tepat Waktu=8, Tepat Waktu=145}
3. Program Studi = S1 Manajemen

- | SKS Total > 95.500
- | | IPK Sem 5 > 2.505
- | | | IPK Sem 5 > 2.930
- | | | | Jenis Kelamin = L
- | | | | | IPK Sem 5 > 2.965
- | | | | | IPK Sem 5 > 3.020: Tepat Waktu {Tidak Tepat Waktu=16, Tepat Waktu=67}
- | | | | | IPK Sem 5 ≤ 3.020: Tidak Tepat Waktu {Tidak Tepat Waktu=2, Tepat Waktu=0}
- | | | | | IPK Sem 5 ≤ 2.965: Tepat Waktu {Tidak Tepat Waktu=0, Tepat Waktu=4}
- | | | | | Jenis Kelamin = P
- | | | | | Asal Mahasiswa = Luar Pringsewu
- | | | | | IPK Sem 5 > 3.470
- | | | | | | IPK Sem 5 > 3.690
- | | | | | | | IPK Sem 5 > 3.715: Tepat Waktu {Tidak Tepat Waktu=0, Tepat Waktu=3}
- | | | | | | | IPK Sem 5 ≤ 3.715: Tidak Tepat Waktu {Tidak Tepat Waktu=1, Tepat Waktu=1}
- | | | | | | | IPK Sem 5 ≤ 3.690: Tepat Waktu {Tidak Tepat Waktu=0, Tepat Waktu=18}
- | | | | | | | IPK Sem 5 ≤ 3.470
- | | | | | | | IPK Sem 5 > 3.440: Tidak Tepat Waktu {Tidak Tepat Waktu=3, Tepat Waktu=0}
- | | | | | | | IPK Sem 5 ≤ 3.440: Tepat Waktu {Tidak Tepat Waktu=1, Tepat Waktu=11}
- | | | | | | | Asal Mahasiswa = Pringsewu
- | | | | | | | IPK Sem 5 > 3.240: Tepat Waktu {Tidak Tepat Waktu=0, Tepat Waktu=83}
- | | | | | | | IPK Sem 5 ≤ 3.240
- | | | | | | | IPK Sem 5 > 3.190: Tidak Tepat Waktu {Tidak Tepat Waktu=1, Tepat Waktu=1}
- | | | | | | | IPK Sem 5 ≤ 3.190: Tepat Waktu {Tidak Tepat Waktu=0, Tepat Waktu=5}
- | | | | | IPK Sem 5 ≤ 2.930
- | | | | | IPK Sem 5 > 2.725: Tidak Tepat Waktu {Tidak Tepat Waktu=8, Tepat Waktu=1}
- | | | | | IPK Sem 5 ≤ 2.725: Tepat Waktu {Tidak Tepat Waktu=0, Tepat Waktu=4}
- | | | IPK Sem 5 ≤ 2.505: Tidak Tepat Waktu {Tidak Tepat Waktu=3, Tepat Waktu=0}
- | | SKS Total ≤ 95.500: Tidak Tepat Waktu {Tidak Tepat Waktu=41, Tepat Waktu=1}
- 4. Program Studi = S1 Pendidikan Bahasa Inggris
 - | | IPK Sem 5 > 2.905
 - | | | SKS Total > 107
 - | | | | IPK Sem 5 > 3.325: Tepat Waktu {Tidak Tepat Waktu=1, Tepat Waktu=20}
 - | | | | IPK Sem 5 ≤ 3.325: Tidak Tepat Waktu {Tidak Tepat Waktu=1, Tepat Waktu=1}
 - | | | SKS Total ≤ 107
 - | | | | SKS Total > 98.500
 - | | | | | IPK Sem 5 > 3.015
 - | | | | | | IPK Sem 5 > 3.095: Tidak Tepat Waktu {Tidak Tepat Waktu=22, Tepat Waktu=16}
 - | | | | | | IPK Sem 5 ≤ 3.095: Tepat Waktu {Tidak Tepat Waktu=0, Tepat Waktu=3}
 - | | | | | | IPK Sem 5 ≤ 3.015: Tidak Tepat Waktu {Tidak Tepat Waktu=3, Tepat Waktu=0}
 - | | | | SKS Total ≤ 98.500: Tepat Waktu {Tidak Tepat Waktu=0, Tepat Waktu=2}
 - | | | IPK Sem 5 ≤ 2.905: Tidak Tepat Waktu {Tidak Tepat Waktu=4, Tepat Waktu=0}
- 5. Program Studi = S1 Pendidikan Bahasa dan Sastra Indonesia
 - | | SKS Total > 105: Tepat Waktu {Tidak Tepat Waktu=1, Tepat Waktu=76}
 - | | SKS Total ≤ 105: Tidak Tepat Waktu {Tidak Tepat Waktu=5, Tepat Waktu=0}
- 6. Program Studi = S1 Pendidikan Guru Sekolah Dasar
 - | | IPK Sem 5 > 3.010

- | | IPK Sem 5 > 3.895: Tepat Waktu {Tidak Tepat Waktu=0, Tepat Waktu=2}
- | | IPK Sem 5 ≤ 3.895: Tidak Tepat Waktu {Tidak Tepat Waktu=82, Tepat Waktu=73}
- | | IPK Sem 5 ≤ 3.010: Tidak Tepat Waktu {Tidak Tepat Waktu=21, Tepat Waktu=0}
- 7. Program Studi = S1 Pendidikan Matematika
 - | IPK Sem 5 > 2.965
 - | | IPK Sem 5 > 3.620: Tepat Waktu {Tidak Tepat Waktu=0, Tepat Waktu=10}
 - | | IPK Sem 5 ≤ 3.620
 - | | | Jenis Kelamin = L: Tidak Tepat Waktu {Tidak Tepat Waktu=7, Tepat Waktu=2}
 - | | | Jenis Kelamin = P
 - | | | | IPK Sem 5 > 3.130
 - | | | | | IPK Sem 5 > 3.170
 - | | | | | | IPK Sem 5 > 3.195
 - | | | | | | Asal Mahasiswa = Luar Pringsewu: Tepat Waktu {Tidak Tepat Waktu=4, Tepat Waktu=14}
 - | | | | | | Asal Mahasiswa = Pringsewu
 - | | | | | | | SKS Total > 110: Tepat Waktu {Tidak Tepat Waktu=7, Tepat Waktu=12}
 - | | | | | | | SKS Total ≤ 110: Tidak Tepat Waktu {Tidak Tepat Waktu=2, Tepat Waktu=1}
 - | | | | | | | IPK Sem 5 ≤ 3.195: Tepat Waktu {Tidak Tepat Waktu=0, Tepat Waktu=2}
 - | | | | | | | IPK Sem 5 ≤ 3.170: Tidak Tepat Waktu {Tidak Tepat Waktu=2, Tepat Waktu=1}
 - | | | | | | | IPK Sem 5 ≤ 3.130: Tepat Waktu {Tidak Tepat Waktu=0, Tepat Waktu=2}
 - | | | | | | | IPK Sem 5 ≤ 2.965: Tidak Tepat Waktu {Tidak Tepat Waktu=6, Tepat Waktu=0}

Berdasarkan hasil pengujian yang sudah dilakukan maka model tersebut layak digunakan sebagai model prediksi kelulusan tepat waktu mahasiswa, kelayakan model yang didapatkan didukung dengan tingkat akurasi dari kedua model yang digunakan dalam penelitian ini yang ditunjukkan pada tabel 4.2 dibawah ini.

Tabel 4.2 Hasil Analisis Tingkat Akurasi Model

Metode	Accuracy	Precision	Recall	AUC
C4.5	79.91%	89.06%	81.38%	0.823
Naive Bayes	76.95%	75.95%	98.38%	0.838

Berdasarkan tabel 4.2 dapat dilihat bahwa metode Algoritma C4.5 lebih unggul dari *Naive Bayes*. Hasil analisis model algoritma C4.5 memiliki tingkat akurasi 79,91 %, nilai AUC 0.823, tingkat *precision* 89,06% dan *recall* 81.38%. Sedangkan *Naive Bayes* memiliki tingkat akurasi 76,95%, nilai AUC 0.838, *tingkat precision* 75.95% dan *recall* 98.38%. Tingginya tingkat akurasi yang diperoleh metode Algoritma C4.5 dipengaruhi oleh banyaknya data yang digunakan, hal tersebut dapat terjadi saat pengujian. Dapat dikatakan lebih baik dari *Naive Bayes* karena tingginya

tingkat akurasi yang diperoleh metode *Naive Bayes* dipengaruhi oleh banyaknya data yang digunakan, perbandingan jumlah *data training* dan *testing* serta tipe atribut yang digunakan dan jumlah variabel yang digunakan. Algoritma C4.5 setiap akan memulai eksekusi di setiap atributnya dikelompokkan setiap atributnya dengan sangat baik dan memberi perlabelan atau range nilai pada setiap proses eksekusi per atribut nya. Sehingga klasifikasinya dapat dikatakan lebih khusus. Karena dilakukan dengan memperlakukan proses per atributnya. Dibuktikan juga pada saat satu kali proses mencari nilai *entropy* per partisi data, hanya untuk mendapatkan salah satu nilai *gain* terbesar dari setiap atribut.