

**IMPLEMENTASI ALGORITMA C4.5 DALAM PREDIKSI
KELULUSAN MAHASISWA STMIK PALANGKARAYA**

TUGAS AKHIR

Disusun untuk Memenuhi Syarat Kelulusan Program Strata I pada
Sekolah Tinggi Manajemen Informatika dan Komputer
(STMIK) Palangkaraya



OLEH

TETI SUMARNI

NIM C1555201023

PROGRAM STUDI TEKNIK INFORMATIKA

**SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
(STMIK) PALANGKARAYA**

2019

**IMPLEMENTASI ALGORITMA C4.5 DALAM PREDIKSI
KELULUSAN MAHASISWA STMIK PALANGKARAYA**

TUGAS AKHIR

Disusun untuk Memenuhi Syarat Kelulusan Program Strata I pada
Sekolah Tinggi Manajemen Informatika dan Komputer
(STMIK) Palangkaraya

Oleh :

TETI SUMARNI

NIM C1555201023

PROGRAM STUDI TEKNIK INFORMATIKA

**SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
(STMIK) PALANGKARAYA**

2019

LEMBAR PERNYATAAN

Yang bertanda tangan dibawah ini :

Nama Mahasiswa : **TETI SUMARNI**
NIM : **C1555201023**

menyatakan bahwa Tugas Akhir dengan judul :

IMPLEMENTASI ALGORITMA C4.5 DALAM PREDIKSI KELULUSAN MAHASISWA STMIK PALANGKARAYA

adalah hasil karya saya dan bukan merupakan duplikasi sebagian atau seluruhnya dari karya orang lain, kecuali bagian yang sumber informasi dicantumkan.

Pernyataan ini dibuat dengan sebenar-benarnya secara sadar dan bertanggungjawab dan saya bersedia menerima sanksi pembatalan Tugas Akhir apabila terbukti melakukan duplikasi terhadap Tugas Akhir atau karya ilmiah lain yang sudah ada.

Palangka Raya, 25 Juli 2019

Yang Membuat Pernyataan,



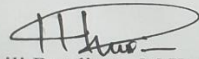
TETI SUMARNI

PERSETUJUAN

**IMPLEMENTASI ALGORITMA C4.5 DALAM PREDIKSI
KELULUSAN MAHASISWA STMIK PALANGKARAYA**

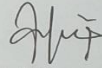
Tugas Akhir ini telah disetujui untuk diujikan
pada Tanggal 23 Juli 2019

Pembimbing I,



Lili Rusdiana, M.Kom
NIK. 198707282011007

Pembimbing II,



Elok Faiqotul Himmah, S.Si., M.Sc
NIK. 198503092009003

Mengetahui

Ketua STMIK Palangkaraya



Suparno, M.Kom.
NIK. 196901041995105

PENGESAHAN

IMPLEMENTASI ALGORITMA C4.5 DALAM PREDIKSI KELULUSAN MAHASISWA STMIK PALANGKARAYA

Tugas Akhir ini telah Diuji, Dinilai dan Disahkan
Oleh Tim Penguji Pada Tanggal 25 Juli 2019

Tim Penguji Tugas Akhir :

1. Frengklin Matatula, S.Kom., M.MSi
Ketua
2. Arliyana, M.Kom
Sekretaris
3. Sam'ani, ST., M.Kom
Anggota
4. Lili Rusdiana, M.Kom
Anggota
5. Elok Faiqotul Himmah, S.Si., M.Sc
Anggota



MOTTO DAN PERSEMBAHAN

*Kuatkanlah hatimu,
janganlah lemah semangatmu
karena ada upah bagi usahamu.*

Tugas Akhir ini kupersembahkan untuk :

- *Kakek dan Nenek tercinta
yang selalu setia mendoakan saya.*
- *Kedua orang tua ku
Yang selalu memberikan dukungan,
doa dan kasih sayang yang teramat
besar.*

ABSTRAK

Teti Sumarni, C1555201023, 2019. *Implementasi Algoritma C4.5 dalam Prediksi Kelulusan Mahasiswa STMIK Palangkaraya*, Pembimbing I Lili Rusdiana, M.Kom., Pembimbing II Elok Faiqotul Himmah, S.Si., M.Sc.

STMIK Palangkaraya telah menyimpan data-datanya kedalam *database* yang terkomputerisasi. Data-data tersebut jika dimanfaatkan dengan maksimal dapat memberikan informasi yang berguna, seperti prediksi kelulusan mahasiswa. Penelitian bertujuan untuk mengimplemetasikan algoritma C4.5 dalam prediksi kelulusan mahasiswa. Penelitian dilakukan di Prodi TI. Data yang digunakan sebagai data latih yaitu data mahasiswa yang telah lulus angkatan 2012 dan 2013. Data yang digunakan sebagai data uji yaitu data mahasiswa angkatan 2014. Atribut yang digunakan adalah jenis kelamin, IPS 1 sampai dengan IPS 6, status pernikahan, status pekerjaan dan jumlah SKS. *Output* sistem yaitu tepat waktu, terlambat dan *drop out*. Implementasi menggunakan Matlab 2016.

Metode yang digunakan adalah *data mining* dengan algoritma C4.5. Metode pengembangan perangkat lunak yang digunakan yaitu *Extreme Programming*. Metode pengumpulan data yang digunakan yaitu wawancara dan studi *literature*.

Berdasarkan uji coba dengan data latih 89 data dan data uji 31 data diperoleh akurasi sebesar 74,19% sedangkan dengan data uji 20 data diperoleh akurasi sebesar 70%. Uji coba dengan data latih 60 data dan data uji 31 data diperoleh akurasi sebesar 38,70% sedangkan dengan data uji 20 data diperoleh akurasi sebesar 35%. Berdasarkan hasil uji coba, semakin banyak data yang digunakan maka semakin besar tingkat akurasi algoritma.

Kata kunci : Algoritma C4.5, *Data Mining*, Prediksi Kelulusan

ABSTRACT

Teti Sumarni, C1555201023, 2019. *Implementasi Algoritma C4.5 dalam Prediksi Kelulusan Mahasiswa STMIK Palangkaraya*, Advisor I Lili Rusdiana, M.Kom., Advisor II Elok Faiqotul Himmah, S.Si., M.Sc.

STMIK Palangkaraya has stored its data into a computerized database. If the data is maximally utilized these data can provide useful information, such as predictions of student graduation. This research aims to implement the C4.5 algorithm in predicting student graduation. This research was conducted in the Prodi TI. The data used as training data is data on students who have graduated in the class of 2012 and 2013. The data used as test data is 2014 student data. The attributes used are gender, IPS 1 to IPS 6, marital status, employments status, and number of SKS. System output which is on time, late and drop out. Implementation using Matlab 2016.

The method used is data mining with algorithm C4.5. The software development method used is Extreme Programming. Data collection methods used are interview and literature study.

Based on trial with training data 89 and 31 test data obtained an accuracy of 74,19% while with 20 test data obtained an accuracy of 70%. Trial with training data 60 and 31 test data obtained an accuracy of 38,71%, while with 20 test data obtained an accuracy of 35%. Based on the results of the trial, the more data used, the greater the level of accuracy of the algorithm.

Kata kunci : C4.5 Algorithm, Data mining, Prediction of Graduation

KATA PENGANTAR

Dengan memanjatkan puji dan syukur kepada Tuhan yang Maha Esa, atas berkat dan rahmatnya sehingga penulis dapat menyelesaikan Tugas Akhir dengan judul **“IMPLEMENTASI ALGORITMA C4.5 DALAM PREDIKSI KELULUSAN MAHASISWA STMIK PALANGKARAYA”**.

Pada kesempatan ini penulis menyampaikan banyak terima kasih yang sebesar-besarnya kepada semua pihak yang telah memberikan bimbingan dan motivasi dalam penyelesaian penulisan Tugas Akhir ini, yaitu kepada :

1. Suparno, M.Kom selaku Ketus Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) Palangkaraya.
2. Lili Rusdiana, M.Kom selaku dosen pembimbing I yang banyak memberikan saran dalam penyelesaian sistem Tugas Akhir ini.
3. Elok Faiqotul Himmah, S.Si., M.Sc selaku dosen pembimbing II yang banyak memberikan koreksi dan saran dalam penyelesaian penulisan Tugas Akhir ini.
4. Kedua orang tua tercinta yang telah memberika dukungan baik moril maupun materi.

Tugas akhir ini memiliki banyak kekurangan oleh karena itu, penulis mengharapkan kritik dan saran yang mendukung untuk penyempurnaan penulisan karya ilmiah.

Palangkaraya, 25 Juli 2019

Penulis

DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN PERNYATAAN	ii
HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN	iv
HALAMAN MOTTO DAN PERSEMBAHAN	v
HALAMAN ABSTRAK	vi
HALAMAN ABSTRACT	vii
KATA PENGANTAR	viii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
DAFTAR PERSAMAAN	xiii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN	
A. Latar Belakang Masalah	1
B. Perumusan Masalah	2
C. Batasan Masalah	2
D. Tujuan dan Manfaat	3
E. Metode Penelitian	4
1. Metode Pengumpulan Data	4
2. Tahap Pengembangan Perangkat Lunak	4
F. Sistematika Penulisan	6
BAB II LANDASAN TEORI	
A. Tinjauan Pustaka	8
B. Kajian Teori	10
1. Data Mining	10
2. Pohon Keputusan (<i>Decision Tree</i>)	13
3. Algoritma C4.5	14
4. Prediksi	20
5. <i>Flowchart</i>	20
6. <i>Black Box Testing</i>	22
C. Perangkat Lunak yang digunakan	24
1. Matlab	22
2. <i>Microsoft Office Excel</i>	26
BAB III ANALISIS DESAIN SISTEM	
A. Tinjauan Umum	27
B. Analisis	28
1. Analisis Kelemahan Sistem	28
2. Analisis Kebutuhan Sistem	28
a. Kebutuhan Perangkat Keras	28
b. Kebutuhan Perangkat Lunak	29
c. Kebutuhan Informasi	29
d. Kebutuhan Pengguna	29
3. Analisis Kelayakan Sistem	29

a. Kelayakan Teknologi	29
b. Kelayakan Hukum.....	30
c. Kelayakan Operasi	30
C. Desain Sistem	31
1. Desain Proses	31
2. Desain <i>interface</i> /Antarmuka	35
BAB IV DESAIN ANALISIS SISTEM	
A. Implementasi	39
1. Uji Coba Sistem dan Program.....	39
2. Manual Program.....	43
B. Pembahasan	50
1. Pembahasan Listing Program.....	50
2. Pembahasan Hasil Penelitian	55
BAB V KESIMPULAN DAN SARAN	
A. Kesimpulan	84
B. Saran.....	84

DAFTAR PUSTAKA
LAMPIRAN

DAFTAR TABEL

Tabel 1. <i>Learning Dataset</i>	15
Tabel 2. Analisis Atribut, <i>Entropy</i> dan <i>Gain</i>	16
Tabel 3. Data yang Memiliki Kelembaban = Tinggi	17
Tabel 4. Hasil Analisis Node 1.1	18
Tabel 5. Data yang Memiliki Cuaca = Hujan	19
Tabel 6. Hasil Analisis Node 1.1.2	19
Tabel 7. Simbol-Simbol <i>Flowchart</i>	21
Tabel 8. Rencana Pengujian Menu <i>Training</i>	39
Tabel 9. Hasil Pengujian Menu <i>Training</i>	40
Tabel 10. Rencana Pengujian Menu <i>Testing</i> 1.....	41
Tabel 11. Hasil Pengujian Menu <i>Training</i> 1.....	41
Tabel 12. Rencana Pengujian Menu <i>Testing</i> 2.....	42
Tabel 13. Hasil Pengujian Menu <i>Testing</i> 2	42
Tabel 14. Data Latih	55
Tabel 15. Transformasi data	56
Tabel 16. Nilai <i>threshold</i> Atribut Jenis Kelamin	57
Tabel 17. Nilai <i>Threshold</i> Atribut IPS 1	57
Tabel 18. Nilai <i>Threshold</i> Atribut IPS 2	58
Tabel 19. Nilai <i>Threshold</i> Atribut IPS 3	58
Tabel 20. Nilai <i>Threshold</i> Atribut IPS 4	59
Tabel 21. Nilai <i>Threshold</i> Atribut IPS 5	59
Tabel 22. Nilai <i>Threshold</i> Atribut IPS 6	60
Tabel 23. Nilai <i>Threshold</i> Atribut Status Pernikahan	60
Tabel 24. Nilai <i>Threshold</i> Atribut Status Pekerjaan	60
Tabel 25. Nilai <i>Threshold</i> Atribut Jumlah SKS	61
Tabel 26. Nilai <i>Threshold</i> Berdasarkan <i>GainRatio</i> Terbesar	72
Tabel 27. Perhitungan Algoritma C4.5	72
Tabel 28. Perhitungan <i>Entropy</i>	74
Tabel 29. Hasil Perhitungan <i>Gain</i>	75
Tabel 30. Data yang Memiliki IPS 1 $\geq 2,62$	79
Tabel 31. Hasil Perhitungan Node 1.1	80
Tabel 32. Perbandingan Pengujian Data	82
Tabel 33. Persentase Perbandingan Data	83

DAFTAR GAMBAR

Gambar 1. Pohon Keputusan Node 1 (<i>root node</i>)	17
Gambar 2. Pohon Keputusan Node 1.1	18
Gambar 3. Pohon Keputusan Akhir	20
Gambar 4. <i>GUIDE Qoict Start</i>	25
Gambar 5. Lembar Kerja Matlab	25
Gambar 6. Lembar Kerja <i>Microsoft Excel 2010</i>	26
Gambar 7. <i>Flowchart Menu Training</i>	31
Gambar 8. <i>Flowchart Menu Testing 1</i>	33
Gambar 9. <i>Flowchart Menu Testing 2</i>	34
Gambar 10. Desain <i>Beranda</i>	35
Gambar 11. Desain <i>Menu Training</i>	36
Gambar 12. Desain <i>Testing 1</i>	37
Gambar 13. Desain <i>Testing 2</i>	37
Gambar 14. Tampilan <i>Beranda</i>	43
Gambar 15. Tampilan <i>Menu Training 1</i>	44
Gambar 16. Tampilan <i>Pencarian File</i>	44
Gambar 17. Tampilan Hasil <i>Load Data</i>	45
Gambar 18. Tampilan Hasil <i>Prediksi</i>	45
Gambar 19. Tampilan <i>Pohon Keputusan</i>	46
Gambar 20. Hasil Akurasi <i>Algoritma C4.5</i>	46
Gambar 21. Tampilan <i>Simpan Tree</i>	47
Gambar 22. Tampilan <i>Menu Testing</i>	47
Gambar 23. Tampilan <i>Load Tree</i>	48
Gambar 24. Tampilan <i>Menu Testing 2</i>	48
Gambar 25. Tampilan Hasil <i>Prediksi</i>	49
Gambar 26. Tampilan <i>Kotak Dialog</i>	50
Gambar 27. <i>Pohon Keputusan Node 1</i>	77
Gambar 28. <i>Pohon Keputusan Node 1 Setelah Dipangkas</i>	79
Gambar 29. <i>Pohon Keputusan Akhir</i>	81
Gambar 30. Hasil <i>Prediksi Sistem</i>	82

DAFTAR PERSAMAAN

Persamaan 1. <i>Gain</i>	14
Persamaan 2. <i>Entropy</i>	15
Persamaan 3. <i>Info (T)</i>	56
Persamaan 4. <i>Info x (T)</i>	63
Persamaan 5. <i>SplitInfo</i>	63
Persamaan 6. <i>Gain (x)</i>	64
Persamaan 7. <i>GainRatio</i>	65
Persamaan 8. Pemangkasan	78

DAFTAR LAMPIRAN

- Lampiran 1. Surat-Surat
- Lampiran 2. Data Latih dan Data Uji
- Lampiran 3. *Black Box Testing*
- Lampiran 4. Tabel Distribusi Normal

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

STMIK Palangkaraya berdiri sesuai dengan ijin yang dikeluarkan oleh Dirjen Dikti Depdikbud Nomor 078/D/O/1995, tanggal 28 September 1995 dan ijin perubahan status menjadi STMIK Palangkaraya dari Menteri Pendidikan Nasional Republik Indonesia Nomor 71/D/O/2007 tanggal 24 Mei 2007. Diera digital saat ini, STMIK Palangkaraya telah menyimpan data-datanya kedalam *database* yang terkomputerisasi seperti data mahasiswa, data dosen, data mahasiswa yang telah lulus, dan data lainnya yang berhubungan dengan STMIK Palangkaraya. Data-data tersebut digunakan untuk keperluan akreditasi dan keperluan lainnya yang berhubungan dengan STMIK Palangkaraya.

Data-data tersebut jika dimanfaatkan dengan maksimal dapat memberikan informasi yang berguna bagi STMIK Palangkaraya. Seperti data mahasiswa yang telah lulus dapat diolah menggunakan *data mining* sehingga menghasilkan suatu informasi yang berguna, seperti prediksi kelulusan mahasiswa. *Data mining* adalah proses untuk mendapatkan informasi yang berguna dari gudang basis data yang besar. Data-data mahasiswa yang telah lulus seperti NIM, nama, IP semester, IPK, dan lain-lain diolah menggunakan *data mining* yang selanjutnya digunakan sebagai prediksi untuk mahasiswa yang belum lulus. Salah satu metode dalam *data mining* yang banyak digunakan yaitu Pohon Keputusan (*Decision Tree*)

dengan algoritma C4.5. Algoritma C4.5 merupakan algoritma klasifikasi pohon keputusan yang banyak digunakan karena memiliki kelebihan utama dari algoritma lainnya diantaranya, dapat menghasilkan pohon keputusan yang mudah diinterpretasikan, memiliki tingkat akurasi yang dapat diterima, serta efisien dalam menangani atribut bertipe diskret dan numerik.

Berdasarkan uraian latar belakang diatas, penulis tertarik untuk mengangkat Tugas Akhir dengan judul **“Implementasi Algoritma C4.5 dalam Prediksi Kelulusan Mahasiswa STMIK Palangkaraya”** diharapkan dengan adanya implementasi algoritma C4.5 ini dapat diketahui mahasiswa yang lulus tepat waktu, tidak tepat waktu, dan *dropout*.

B. Perumusan Masalah

Berdasarkan latar belakang masalah, penulis dapat merumuskan masalah yang akan diselesaikan yaitu bagaimana mengimplementasikan algoritma C4.5 dalam prediksi kelulusan mahasiswa STMIK Palangkaraya?

C. Batasan Masalah

Agar pembahasan tidak melebar luas keluar dari judul penulisan, maka penulis akan melakukan pembatasan masalah yaitu sebagai berikut :

1. Penelitian dilakukan di STMIK Palangkaraya, Program Studi Teknik Informatika. Data yang akan digunakan sebagai data latih yaitu data mahasiswa yang telah lulus angkatan tahun 2012 sampai dengan tahun

2013 dan data yang akan digunakan sebagai data uji yaitu data mahasiswa angkatan tahun 2014,

2. Atribut yang akan digunakan yaitu IP semester 1 sampai dengan semester 6, status pernikahan, status pekerjaan, jenis kelamin dan jumlah SKS yang telah dikumpulkan,
3. Keluaran sistem berupa hasil prediksi yaitu lulus tepat waktu, tidak tepat waktu dan *drop out*,
4. Bahasa pemrograman yang akan digunakan untuk implementasi yaitu Matlab.

D. Tujuan dan Manfaat

1. Tujuan

Adapun tujuan dari penulis adalah mengimplementasikan algoritma C4.5 dalam prediksi kelulusan mahasiswa STMIK Palangkaraya untuk mengetahui mahasiswa yang lulus tepat waktu, tidak tepat waktu, dan *drop out*.

2. Manfaat

Penelitian ini diharapkan dapat memberi manfaat, yaitu :

a. Bagi Penulis

Manfaat penelitian ini bagi penulis adalah untuk mengimplementasikan ilmu yang telah didapat dibangku perkuliahan dan sebagai syarat kelulusan jenjang Strata I.

b. Bagi STMIK Palangkaraya

Manfaat penelitian ini bagi STMIK Palangkaraya adalah menambah referensi pada perpustakaan.

c. Bagi Program Studi Teknik Informatika

Manfaat penelitian ini bagi Program Studi Teknik Informatika adalah mengetahui mahasiswa yang lulus tepat waktu, terlambat atau *drop out*.

E. Metode Penelitian

1. Tahap Pengumpulan Data

a. Metode Observasi

Observasi merupakan kegiatan pemuatan penelitian terhadap suatu objek . Dalam hal ini penulis melakukan pengumpulan data dengan melakukan pengumpulan data di Kasubag Kemahasiswaan dan Alumni.

b. Studi *Literature*

Studi *literature* adalah mencari referensi teori yang relevan dengan kasus atau permasalahan yang ditemukan. Dalam hal ini penulis melakukan pengumpulan data dengan mencari referensi melalui jurnal dan bacaan yang ada kaitannya dengan topik penelitian.

2. Tahap Pengembangan Perangkat Lunak

Metode pengembangan perangkat lunak yang akan digunakan dalam penelitian ini adalah *Extreme Programming* (XP).

Extreme programming merupakan sebuah proses rekayasa perangkat lunak yang cenderung menggunakan pendekatan berorientasi objek dan sasaran dari metode ini adalah tim yang dibentuk dalam skala kecil sampai medium serta metode ini juga sesuai jika tim dihadapkan dengan *requirement* yang tidak jelas maupun terjadi perubahan-perubahan *requirement* yang sangat cepat (Supriyatna, 2018). Adapun tahapan-tahapannya antara lain (Oktaviani & Hutrianto, 2016) :

a. *Planning*

Tahapan ini dimulai dengan mendengarkan kumpulan kebutuhan aktifitas suatu sistem yang memungkinkan pengguna memahami proses bisnis untuk sistem dan mendapatkan gambaran yang jelas mengenai fitur utama, fungsionalitas, dan keluaran yang diinginkan . Dalam hal ini penulis melakukan analisis kebutuhan sistem, kebutuhan pengguna, dan mengidentifikasi permasalahan yang timbul pada sistem berjalan.

b. *Design*

Pada tahapan ini dilakukan pembuatan model sistem berdasarkan hasil analisis kebutuhan yang didapatkan. Pemodelan sistem yang digunakan yaitu *Flowchart*.

c. *Coding*

Tahapan ini merupakan implementasi dari perancangan model sistem yang telah dibuat kedalam kode program. Bahasa pemrograman yang akan digunakan untuk implementasi yaitu Matlab.

d. *Testing*

Tahapan ini merupakan tahapan pengujian terhadap sistem yang telah dibangun, pada tahapan ini berfokus pada fitur dan fungsionalitas dari keseluruhan sistem kemudian ditinjau oleh pengguna sistem. Metode yang digunakan dalam melakukan pengujian sistem yaitu *Black-Box* dengan melakukan pengujian terhadap masukan dan keluaran yang dihasilkan oleh sistem.

F. Sistematika Penulisan

Untuk memudahkan dalam memahami laporan tugas akhir, penulis mengemukakan sistematika penulisan yaitu sebagai berikut :

BAB I PENDAHULUAN

Bab ini berisi latar belakang masalah yang menjelaskan alasan penulis tertarik untuk mengangkat sebagai judul tugas akhir, perumusan masalah yang akan diselesaikan, batasan masalah, tujuan dan manfaat penelitian, metode penelitian yang digunakan dalam tahap pengumpulan data dan tahap pengembangan perangkat lunak, serta sistematika penulisan laporan tugas akhir.

BAB II LANDASAN TEORI

Bab ini berisi tinjauan pustaka tentang hasil penelitian orang lain dengan penelitian yang sedang dilakukan, kajian teori yang berkaitan dengan judul penelitian serta teori tentang perangkat lunak yang digunakan.

BAB III ANALIS DAN DESAIN SISTEM

Bab ini berisi tentang analisis terhadap permasalahan yang akan diselesaikan seperti tinjauan umum, analisis kelemahan sistem, analisis kebutuhan sistem baik itu perangkat keras, perangkat lunak dan informasi. Analisis kelayakan sistem dari segi teknologi, hukum, dan operasional. Dan desain sistem yang terdiri dari desain proses dan desain antarmuka.

BAB IV IMPLEMENTASI DAN PEMBAHASAN

Bab ini berisi implementasi sistem, hasil penelitian serta pembahasannya.

BAB V KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan yang dapat diambil dari hasil penelitian dan saran-saran dari kelemahan sistem yang telah dibangun agar dapat dikembangkan lebih lanjut.

BAB II

LANDASAN TEORI

A. Tinjauan Pustaka

Untuk membantu dalam proses penelitian ini, penulis melakukan penelusuran untuk membandingkan antara penelitian yang dilakukan penulis dengan penelitian yang dilakukan oleh pihak lain. Diantaranya sebagai berikut :

1. Penelitian yang dilakukan oleh Fitriana Harahap (2015) dengan judul “Penerapan Data Mining dalam Memprediksi Pembelian Cat”. Penelitian ini menggunakan atribut seperti kualitas, animo, harga, dan kompetensi. Berdasarkan hasil pengujian diketahui bahwa *data mining* dengan algoritma C4.5 bermanfaat sekali dalam proses pengambilan keputusan dalam pembelian cat pada *Home Smart* Medan. Faktor tertinggi yang mempengaruhi pembelian cat adalah faktor kompetensi, faktor kedua yaitu kualitas dan animo, sedangkan faktor harga tidak mempengaruhi pembelian cat karena cat dengan harga mahal ternyata diminati oleh pelanggan,
2. Penelitian yang dilakukan oleh Triowali Rosandy (2016) dengan judul “Perbandingan Metode *Naive Bayes Classifier* dengan Metode *Decision Tree* (C4.5) untuk Menganalisis Kelancaran Pembiayaan (Study Kasus : KSPPS/BMT Al-Fadhila)”. Penelitian ini menggunakan atribut seperti pekerjaan, jumlah, jenis, lama, dan

jaminan. Melalui pengujian yang dilakukan dengan menggunakan pemodelan metode *Naive Bayes Classifier* didapat tingkat akurasi tertinggi sebesar 50% sedangkan pada pemodelan *Decison Tree* (C4.5) memperoleh rata-rata tingkat akurasi tertinggi sebesar 56%. Oleh karena itu berdasarkan tingkat akurasi, metode *Decison Tree* (C4.5) lebih baik dalam melakukan prediksi pembiayaan pada KSPPS/BMT Al-Fadhila,

3. Penelitian yang dilakukan oleh Agus Romadhona, Suprapedi dan H. Himawan (2017) dengan judul “Prediksi Kelulusan Mahasiswa Tepat Waktu Berdasarkan Usia, Jenis Kelamin, dan Indeks Prestasi Menggunakan Algoritma *Decision Tree*”. Melalui pengujian yang dilakukan didapat bahwa kinerja algoritma C4.5 lebih baik dibandingkan dengan algoritma *ID3* dan *Chaid*. Serta dari perhitungan nilai *information gain* pada atribut Indeks Prestasi mendapat nilai *gain* tertinggi yaitu 0,340 ,
4. Penelitian yang dilakukan oleh Bambang Hermanto, Azhari SN, dan Fajri Profesio Putra (2017) dengan judul “Analisis Kinerja *Decision Tree* dalam Prediksi Potensi Pelunasan Kredit Calon Debitur”. Penelitian ini dilakukan untuk mengetahui lama waktu eksekusi (*execution time*) dalam membangun pohon keputusan (*generate tree*) dan aturan keputusan, serta nilai prosentase akurasi data. Melalui pengujian yang dilakukan telah dibangun rancang bangun aplikasi dengan menerapkan algoritma C4.5 dan

berdasarkan 5 kategori yang uji dalam proses *generate tree* dibutuhkan rata-rata waktu 112 detik, *generate rules* dibutuhkan rata-rata waktu 1,78 detik, dan prosentase rata-rata nilai akurasi data 51,2 % ,

5. Penelitian yang dilakukan oleh Nurul Azwati (2018) dengan judul “Analisis Algoritma C4.5 untuk Memprediksi Penjualan Motor pada PT. Capella Dinamik Nusantara Cabang Muka Kuning”. Penelitian ini menggunakan atribut seperti semester, warna, tahun produksi, segmen, harga dan penjualan serta perangkat lunak yang digunakan yaitu *Weka*. Melalui pengujian yang dilakukan algoritma C4.5 dapat digunakan sebagai metode klasifikasi dalam memperdiksi penjualan motor oleh PT. Capella Dinamik Nusantara Cabang Muka Kuning. Atribut tertinggi dalam memperdiksi penjualan adalah atribut penjualan.

B. Kajian Teori

1. Data Mining

Data mining adalah proses yang menggunakan statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *database* besar. *Data mining* adalah analisis pengamatan *dataset* untuk menemukan hubungan tak terduga dan untuk meringkas data dengan cara baru yang baik dimengerti dan berguna untuk pengguna data. Inti tugas

data mining adalah pemodelan prediktif, analisis asosiasi, analisis *cluster* dan deteksi terhadap anomali. Analisis *data mining* berjalan pada data yang cenderung terus membesar untuk mendapatkan kesimpulan dan keputusan paling layak (Romadhona, Suprapedi & Himawan, 2017).

Data mining memiliki beberapa sebutan atau nama lain yaitu, *Knowledge discovery (mining) indatabases* (KDD), ekstraksi pengetahuan (*knowledge extraction*), analisis data/pola, kecerdasan bisnis (*business intelligence*) dan lain-lain. *Knowledge discovery (mining) indatabases* (KDD) yaitu sebuah proses mencari pengetahuan yang bermanfaat dari data. Tahapan dalam *data mining* secara garis besar dimulai dari data sumber dan berakhir dengan adanya informasi yang dihasilkan dari beberapa tahapan. Proses KDD secara garis besar adalah sebagai berikut (Romadhona, Suprapedi & Himawan, 2017) :

a. Seleksi Data

Pemilihan (seleksi) data baru dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang akan digunakan untuk proses *data mining* disimpan dalam suatu berkas terpisah dari basis data operasional.

b. Pembersihan Data

Sebelum proses *data mining* dapat dilaksanakan perlu dilakukan proses pembersihan pada data yang menjadi fokus KDD. Proses pembersihan mencakup data yang *inkonsisten* dan memperbaiki kesalahan pada data seperti kesalahan cetak (*tipografi*).

c. Transformasi

Pada tahap transformasi data diubah kedalam bentuk yang sesuai dengan format dalam *data mining*. Beberapa teknik *data mining* membutuhkan format data yang khusus sebelum diaplikasikan. Sebagai contoh, beberapa teknik standar seperti analisis asosiasi dan klustering hanya bisa menerima input data kategorikal. Disini juga dilakukan pemilihan data yang diperlukan oleh teknik *data mining* yang dipakai.

d. *Data Mining*

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode atau algoritma dalam *data mining* sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan.

e. Interpretasi/Evaluasi

Pola informasi yang dihasilkan dari proses *data mining* perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses KDD yang disebut dengan *interpretation*. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya.

2. Pohon Keputusan (*Decision Tree*)

Pohon keputusan merupakan metode klasifikasi dan prediksi yang sangat terkenal. Metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang merepresentasikan aturan. Aturan dapat dengan mudah dipahami dengan bahasa alami, juga dapat diekspresikan dalam bentuk bahasa basis data seperti *Structured Query Language* untuk mencari *record* pada kategori tertentu. Pohon keputusan juga berguna untuk mengeksplorasi data, menemukan hubungan tersembunyi antara sejumlah calon variabel input dengan sebuah variabel target. Karena pohon keputusan memadukan antara eksplorasi data dan pemodelan, pohon keputusan sangat bagus sebagai langkah awal dalam proses pemodelan bahkan ketika dijadikan sebagai model akhir dari beberapa teknik lain (Hermanto, Azhari & Putra, 2017).

3. Algoritma C4.5

Algoritma C4.5 yaitu sebuah algoritma yang digunakan untuk membangun *decision tree* (pohon keputusan). Algoritma C4.5 merupakan pengembangan dari algoritma ID3. Algoritma C4.5 dan ID3 diciptakan oleh J. Rose Quinlan pada akhir tahun 1970-an. Algoritma C4.5 membuat pohon keputusan dari atas ke bawah, atribut paling atas merupakan akar dan yang paling bawah dinamakan daun (Azwanti, 2018).

Secara umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut (Hermanto, Azhari & Putra , 2017) :

- a. pilih atribut sebagai akar,
- b. buat cabang untuk masing-masing nilai,
- c. bagi kasus dalam cabang,
- d. ulangi proses masing-masing cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Untuk memilih atribut sebagai akar, didasarkan pada nilai *gain* tertinggi dari atribut-atribut yang ada. Untuk menghitung *gain* digunakan rumus seperti Persamaan 1.

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

Keterangan :

S = himpunan kasus

A = atribut

n = jumlah partisi atribut A

|Si| = jumlah kasus pada partisi ke i

|S| = jumlah kasus dalam S

i = 1,2,.....n

Sebelum mendapatkan nilai *gain* yang harus dilakukan yaitu mencari nilai *entropy*. *Entropy* digunakan untuk menentukan seberapa informatif sebuah masukan atribut untuk menghasilkan sebuah atribut. Rumus dasar dari *entropy* seperti pada Persamaan 2.

$$Entropy(S) = \sum_{i=1}^n -p_i * \log_2 p_i \quad (2)$$

Keterangan :

S = himpunan kasus

n = jumlah partisi S

pi = proporsi dari Si terhadap S

i = 1,2,.....n

Contoh algoritma C4.5 untuk menyelesaikan kasus suatu pertandingan tenis akan dilakukan atau tidak, berdasarkan keadaan cuaca, suhu, kelembaban dan angin (www.academia.edu, 2019).

Tabel 1. *Learning Dataset*

No.	Cuaca	Suhu	Kelembaban	Berangin	Main
1	Cerah	Panas	Tinggi	Salah	Tidak
2	Cerah	Panas	Tinggi	Benar	Tidak
3	Berawan	Panas	Tinggi	Salah	Ya
4	Hujan	Sejuk	Tinggi	Salah	Ya
5	Hujan	Dingin	Normal	Salah	Ya
6	Hujan	Dingin	Normal	Benar	Ya
7	Berawan	Dingin	Normal	Benar	Ya
8	Cerah	Sejuk	Tinggi	Salah	Tidak
9	Cerah	Dingin	Normal	Salah	Ya

No.	Cuaca	Suhu	Kelembaban	Berangin	Main
10	Hujan	Sejuk	Normal	Salah	Ya
11	Cerah	Sejuk	Normal	Benar	Ya
12	Berawan	Sejuk	Tinggi	Benar	Ya
13	Berawan	Panas	Normal	Salah	Ya
14	Hujan	Sejuk	Tinggi	Benar	Tidak

Sumber : www.academia.edu, 2019

Langkah pertama hitung *entropy* dari keseluruhan kasus menggunakan Persamaan 2 :

$$\begin{aligned}
 \text{Entropy } (S) &= \left(-\left(\frac{10}{14}\right) \times \log_2 \left(\frac{10}{14}\right)\right) + \left(-\left(\frac{4}{14}\right) \times \log_2 \left(\frac{4}{14}\right)\right) \\
 &= 0.863120569
 \end{aligned}$$

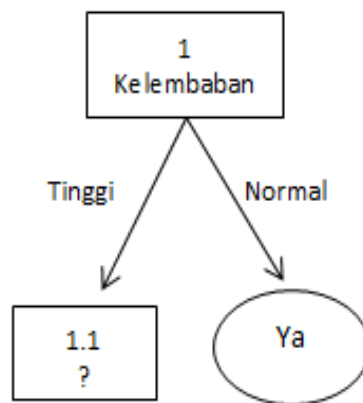
Setelah mendapatkan *entropy* dari keseluruhan kasus, lakukan analisis pada setiap nilai atribut dengan menghitung *entropy* masing-masing nilai atribut menggunakan Persamaan 2. Setelah mendapatkan nilai *entropy* dari masing-masing atribut selanjutnya hitung nilai *gain* menggunakan Persamaan 1.

Tabel 2. Analisis Atribut, *Entropy* dan *Gain*

Node	Atribut	Nilai	Sum (Nilai)	Sum (Ya)	Sum (Tidak)	<i>Entropy</i>	<i>Gain</i>
1		Total	14	10	4	0.863120569	
	Cuaca	Berawan	4	4	0	0	0.258521037
		Hujan	5	4	1	0.721928095	
		Cerah	5	2	3	0.970950594	
	Suhu	Dingin	4	4	0	0	0.183850925
		Sejuk	4	2	2	1	
		Panas	6	4	2	0.918295834	
	Kelembaban	Tinggi	7	3	4	0.985228136	0.370506501
		Normal	7	7	0	0	
	Berangin	Salah	8	6	2	0.811278124	0.005977711
		Benar	6	2	4	0.918295834	

Sumber : www.academia.edu, 2019

Karena nilai *gain* terbesar adalah *gain* kelembaban dengan nilai 0,370506501 , maka kelembaban menjadi node akar (*root node*). Kelembaban normal memiliki 7 kasus dan semuanya memiliki jawaban Ya ($\text{Sum}(\text{Total}) / \text{Sum}(\text{Ya}) = 7/7 = 1$). Dengan demikian, kelembaban normal menjadi daun atau *leaf*.



Gambar 1. Pohon Keputusan Node 1 (*root node*)
Sumber : www.academia.edu, 2019

Selanjutnya node 1.1 akan dianalisis dengan mengambil data yang memiliki kelembaban = tinggi

Tabel 3. Data yang Memiliki Kelembaban = Tinggi

No.	Cuaca	Suhu	Kelembaban	Berangin	Main
1	Cerah	Panas	Tinggi	Salah	Tidak
2	Cerah	Panas	Tinggi	Benar	Tidak
3	Berawan	Panas	Tinggi	Salah	Ya
4	Hujan	Sejuk	Tinggi	Salah	Ya
5	Cerah	Sejuk	Tinggi	Salah	Tidak
6	Berawan	Sejuk	Tinggi	Benar	Ya
7	Hujan	Sejuk	Tinggi	Salah	Tidak

Sumber : www.academia.edu, 2019

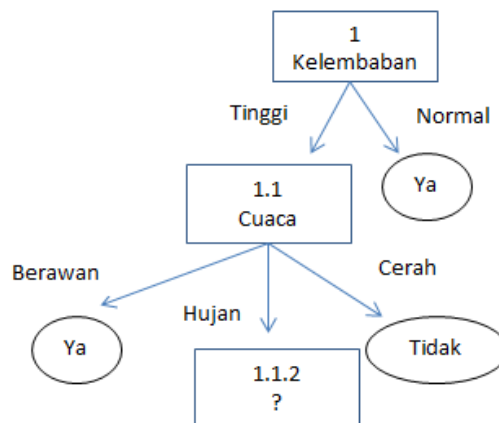
Kemudian analisis dan hitung *entropy* atribut kelembaban tinggi dan *entropy* dari masing-masing atribut serta *gain*nya.

Tabel 4. Hasil Analisis Node 1.1

Node	Atribut	Nilai	Sum (Nilai)	Sum (Ya)	Sum (Tidak)	Entropy	Gain
1.1		Kelembaban = Tinggi	7	3	4	0.985228136	
	Cuaca	Berawan	2	2	0	0	0.69951385
		Hujan	2	1	1	1	
		Cerah	3	0	3	0	
	Suhu	Dingin	0	0	0	0	0.020244207
		Sejuk	3	1	2	0.918295834	
		Panas	4	2	2	1	
	Berangin	Salah	4	2	2	1	0.020244207
		Benar	3	2	1	0.918295834	

Sumber : www.academia.edu, 2019

Dari Tabel 4 nilai *gain* tertinggi ada pada atribut cuaca yaitu 0.69951385 dan cuaca memiliki kasus berjumlah 7 yang terdiri dari berawan 2 kasus dengan nilai ya semua, cerah 3 kasus dengan nilai tidak semua, dan hujan 2 kasus dengan nilai ya dan tidak. Sehingga berawan dan cerah dijadikan sebagai daun atau *leaf*.



Gambar 2. Pohon Keputusan Node 1.1

Sumber : www.academia.edu, 2019

Untuk menganalisis node 1.1.2 lakukan langkah-langkah yang sama seperti sebelumnya.

Tabel 5. Data yang Memiliki Cuaca= Hujan

No.	Cuaca	Suhu	Kelembaban	Berangin	Main
1	Hujan	Sejuk	Tinggi	Salah	Ya
2	Hujan	Sejuk	Tinggi	Benar	Tidak

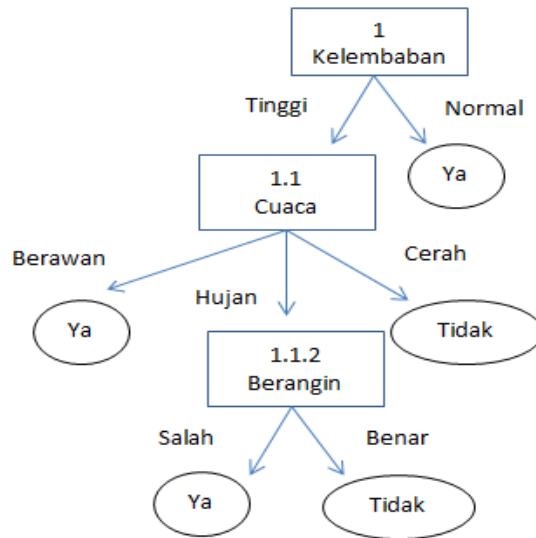
Sumber : www.academia.edu, 2019

Tabel 6. Hasil Analisis Node 1.1.2

Node	Atribut	Nilai	Sum (Nilai)	Sum (Ya)	Sum (Tidak)	Entropy	Gain
1.1.2		Cuaca = Hujan	2	1	1	1	
	Suhu	Dingin	0	0	0	0	0
		Sejuk	0	0	0	0	
		Panas	2	1	1	1	
	Berangin	Salah	1	1	0	0	1
		Benar	1	0	1	0	

Sumber : www.academia.edu, 2019

Tabel 6 menunjukan nilai *gain* tertinggi ada pada atribut berangin yaitu 1 dan berangin memiliki kasus berjumlah 2 yang terdiri dari salah 1 kasus dengan nilai ya dan benar 1 kasus bernilai tidak. Sehingga nilai salah dan benar dijadikan sebagai daun atau *leaf*.



Gambar 3. Pohon Keputusan Akhir
Sumber : www.academia.edu, 2019

Berdasarkan node 1.1.2 semua kasus sudah mempunyai kelas sehingga pohon keputusan diatas merupakan pohon keputusan terakhir yang terbentuk.

4. Prediksi

Prediksi adalah suatu proses memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi di masa depan berdasarkan informasi masa lalu dan sekarang yang dimilikinya. Prediksi dalam *data mining* hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang.

5. Flowchart


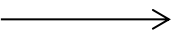

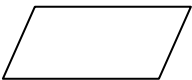
Flowchart menggambarkan urutan logika dari suatu prosedur pemecahan masalah, sehingga *flowchart* merupakan langkah-

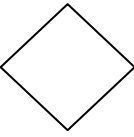
langkah penyelesaian masalah yang dituliskan dalam simbol-simbol tertentu. Diagram alir ini akan menunjukkan alur didalam program secara logika. Diagram alir ini selain dibutuhkan sebagai alat komunikasi, juga diperlukan sebagai dokumentasi (Sitorus, 2015).

Flowchart adalah penggambaran secara grafik dari langkah-langkah dan urutan prosedur dari suatu program. *Flowchart* menolong *analyst* dan *programmer* untuk memecahkan masalah kedalam segmen-segmen yang lebih kecil dan menolong dalam menganalisis alternatif-alternatif lain dalam pengoperasian. *Flowchart* biasanya mempermudah penyelesaian suatu masalah khususnya masalah yang perlu dipelajari dan dievaluasi lebih lanjut (Adelia & Setiawan, 2011).

Berdasarkan pendapat diatas, dapat disimpulkan bahwa *flowchart* merupakan suatu bagan yang dengan simbol-simbol tertentu yang menggambarkan urutan prosedur pemecahan masalah dalam suatu program.

Tabel 7. Simbol-Simbol *Flowchart*

Simbol	Nama	Fungsi
	<i>Terminator</i>	Permulaan/akhir program
	Garis alir (<i>flow line</i>)	Arah aliran program
	Proses	Proses perhitungan/proses pengolahan data
	<i>Input/output data</i>	Proses <i>input/output</i> data,parameter, informasi

Simbol	Nama	Fungsi
	<i>Decision</i>	Perbandingan pernyataan, penyeleksian data yang memberikan pilihan untuk langkah selanjutnya.

Sumber : (Taupik , Irfan & Nurpianti : 2013)

6. *Black Box-Testing*

Black box testing bertujuan untuk menunjukkan fungsi perangkat lunak tentang cara beroperasinya, apakah pemasukan data keluaran telah berjalan sebagaimana yang diharapkan, dan apakah informasi yang disimpan serta *eksternal* selalu dijaga kemukahirannya (Maharani & Merlina, 2014).

Black box testing cenderung untuk menemukan hal-hal berikut (Mustaqbal, Firdaus & Rahmadi, 2015) :

- Fungsi yang tidak benar atau tidak ada;
- Kesalahan antarmuka (*interface errors*);
- Kesalahan pada struktur data dan akses basis data;
- Kesalahan performansi (*performance errors*);
- Kesalahan inisialisasi dan terminas.

C. Perangkat Lunak yang digunakan

1. Matlab

Matlab merupakan perangkat lunak yang digunakan untuk pemrograman, analisis, serta komputasi teknik dan matematis berbasis matriks. Matlab adalah singkatan dari *Matrix Laboratory* karena mampu menyelesaikan masalah perhitungan

dalam bentuk matriks. Matlab versi pertama dirilis pada tahun 1970 oleh Cleve Moler. Pada awalnya, Matlab didesain untuk menyelesaikan masalah-masalah permasalahan aljabar linier. Seiring berjalannya waktu, program ini terus mengalami perkembangan dari segi fungsi dan performa komputasi (Tjolleng, 2017).

Bahasa pemrograman yang kini dikembangkan oleh *MathWorks Inc.* menggabungkan proses pemrograman, komputasi, dan visualisasi melalui lingkungan kerja yang mudah digunakan. Matlab juga memiliki keunggulan umum lainnya, seperti analisis dan eksplorasi data, pengembangan algoritma, pemodelan dan simulasi, visualisasi plot dalam bentuk 2D dan 3D, hingga pengembangan aplikasi antarmuka grafis (Tjolleng, 2017).

Matlab dapat dioperasikan pada sistem operasi *Windows*, *Linux*, maupun *macOS*. Selain itu, Matlab juga bisa dihubungkan dengan aplikasi atau bahasa pemrograman eksternal lainnya, seperti *C*, *Java*, *NET*, dan *Microsoft Excel*. Dalam Matlab tersedia pula *toolbox* yang dapat digunakan untuk aplikasi-aplikasi khusus, seperti pengolahan sinyal, sistem kontrol, logika *fuzzy*, jaringan syaraf tiruan, optimasi, pengolahan citra digital, bioinformatika, simulasi, dan berbagai teknologi lainnya (Tjolleng, 2017).

a. Lingkungan Kerja Matlab

Secara umum, layar utama Matlab memuat tiga bagian penting yaitu sebagai berikut (Tjolleng, 2017) :

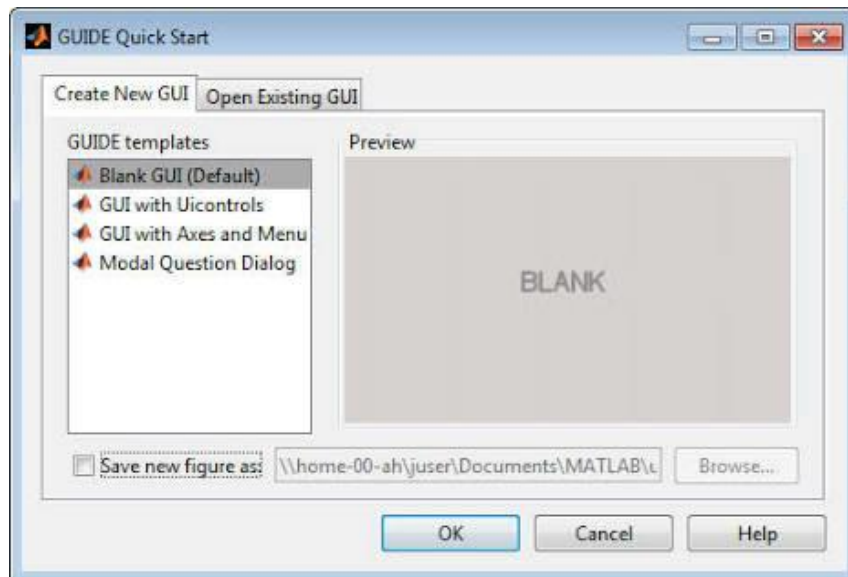
- 1) *Current folder* (folder terkini), untuk mengakses berbagai *file* dalam *direktori* terkini yang sedang digunakan;
- 2) *Command window* (jendela perintah), untuk menuliskan perintah dalam bentuk sintaks program. Perintah yang ditulis dapat berupa perhitungan sederhana, pemanggilan fungsi, demo program, dan lain sebagainya. Setiap penulisan perintah selalu diawali dengan *prompt* “>>”;
- 3) *Workspace* (jendela ruang kerja), untuk mengeksplorasi data yang telah dibuat atau di-*import* dari *file* lain.

b. GUI (*Graphical User Interface*) Matlab

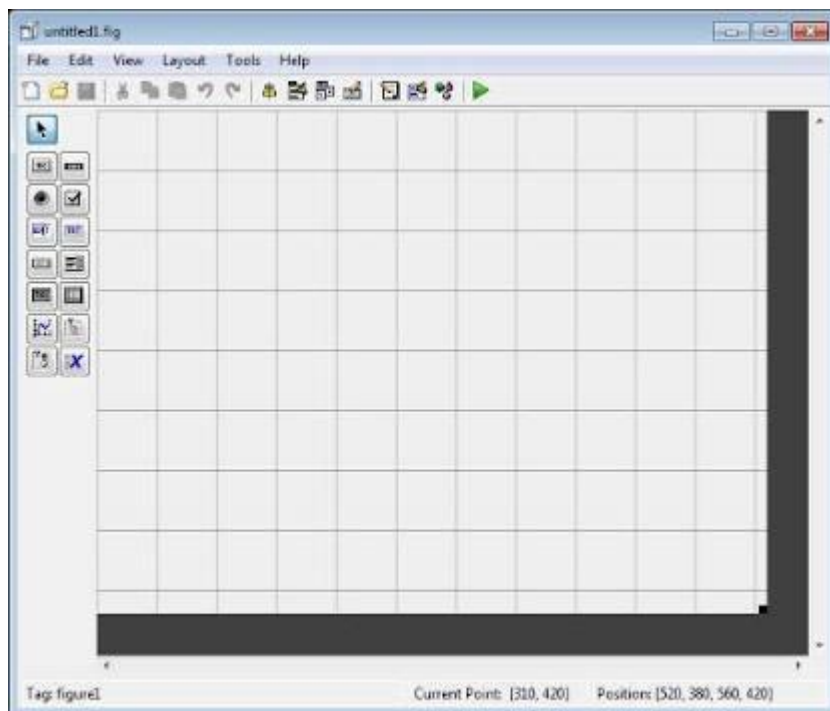
Dalam Matlab tersedia *tool* yang dapat digunakan untuk membuat *graphic user interface* atau antarmuka grafis yang disebut dengan *GUI builder* atau GUIDE (*GUI Development Environment*). GUI menjadi perantara dalam interaksi antara pengguna (*user*) dengan komputer melalui ikon grafis. GUI memudahkan *user* dalam mengoperasikan program yang telah dibuat. Dengan adanya GUI, *user* tidak perlu lagi menjalankan program sendiri melalui jendela perintah (Tjolleng, 2017).

Untuk menjalankan GUI Matlab, terlebih dahulu jalankan program Matlab lalu pada *Command Window* ketik perintah

>>*Guide* maka akan muncul *form* seperti pada gambar 4. Klik *OK* untuk membuat lembar kerja GUI, maka akan muncul lembar kerja seperti pada gambar 5.



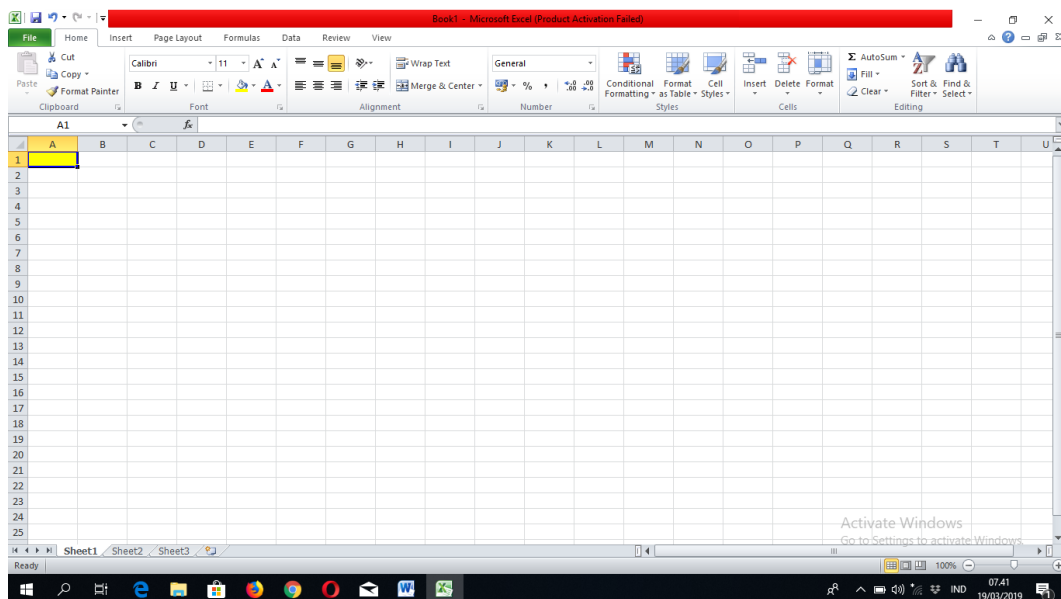
Gambar 4. *GUIDE Quick Start*



Gambar 5. Lembar Kerja GUI Matlab

2. *Microsoft Office Excel*

Microsoft office excel merupakan sebutan sebuah aplikasi lembar kerja *spreadsheet* yang dibuat dan didistribusikan oleh *Microsoft Corporation* untuk sistem operasi *Windows* dan *Mac OS*. Aplikasi ini memiliki fitur kalkulasi dan pembuatan grafik yang dengan menggunakan strategi *marketing Microsoft* yang agresif, menjadikan *Microsoft office excel* sebagai salah satu program komputer yang populer digunakan didalam komputer mikro hingga saat ini (Winahyu, 2016).



Gambar 6. Lembar Kerja *Microsoft Excel* 2010

BAB III

ANALISA DAN DESAIN SISTEM

A. Tinjauan Umum

Algoritma C4.5 merupakan algoritma klasifikasi pohon keputusan yang banyak digunakan karena memiliki kelebihan utama dari algoritma lainnya diantaranya, dapat menghasilkan pohon keputusan yang mudah diinterpretasikan, memiliki tingkat akurasi yang dapat diterima, serta efisien dalam menangani atribut bertipe diskret dan numerik. Dalam penelitian tugas akhir ini, data yang digunakan sebagai data latih yaitu data mahasiswa yang telah lulus angkatan tahun 2011 sampai dengan tahun 2013 dan data yang akan digunakan sebagai data uji yaitu data mahasiswa tahun 2014. Data-data mahasiswa yang telah lulus diolah menggunakan algoritma C4.5 untuk membentuk pohon keputusan yang akan digunakan untuk memprediksi kelulusan mahasiswa yang belum lulus. Atribut yang digunakan yaitu IP semester 1 sampai dengan semester 6, status pernikahan, status pekerjaan, jenis kelamin, dan jumlah SKS yang telah dikumpulkan.

Diharapkan dengan adanya implementasi algoritma C4.5 ini, dapat diketahui mahasiswa yang lulus tepat waktu, tidak tepat waktu, dan *drop out*.

B. Analisis

1. Analisis Kelemahan Sistem

Saat ini di STMIK Palangkaraya belum terdapat sistem untuk prediksi kelulusan mahasiswa berbasis komputer. Program studi hanya menebak kelulusan mahasiswanya berdasarkan pencapaian akademis

yang disajikan pada daftar nilai kumulatif. Akan tetapi data menunjukkan terdapat beberapa mahasiswa dengan pencapaian akademis baik namun status kelulusannya kurang baik.

Berdasarkan permasalahan tersebut peneliti mengusulkan sebuah sistem yang dapat melakukan prediksi kelulusan mahasiswa berdasarkan nilai IP semester 1 sampai dengan semester 6, status pernikahan, status pekerjaan, jenis kelamin dan jumlah SKS yang telah dikumpulkan. Dengan adanya sistem tersebut program studi dapat mengetahui faktor yang dapat mempengaruhi kelulusan mahasiswa, sehingga program studi dapat mengambil kebijakan-kebijakan dalam sistem pembelajaran supaya mahasiswa dapat menyelesaikan studinya tepat waktu.

2. Analisis Kebutuhan Sistem

a. Kebutuhan Perangkat Keras

Pada penelitian ini digunakan Laptop dengan spesifikasi sebagai berikut :

- 1) *Processor* : Intel® Core™ i5-7200U Processor (2.5 GHz, 3M Cache)
- 2) *Memory* : RAM 4GB DDR4
- 3) *GPU* : Intel HD Graphics 620 dan NVIDIA GeForce GT930MX 2GB
- 4) *Ukuran layar* : 14 Inch (1366 x 768 piksel)

b. Kebutuhan Perangkat Lunak

Pada penelitian ini perangkat lunak yang digunakan yaitu sebagai berikut :

- 1) Sistem Operasi *Windows 10*
- 2) Matlab
- 3) *Microsoft Excel*

c. Kebutuhan Informasi

Dalam penelitian ini informasi yang dibutuhkan berupa IP semester 1 sampai dengan semester 6, status pernikahan, status pekerjaan, jenis kelamin, dan jumlah SKS yang telah dikumpulkan. Informasi-informasi tersebut didapat dari hasil pengumpulan data dari Kasubag Kemahasiswaan dan Alumni STMIK Palangkaraya.

d. Kebutuhan Pengguna

Pengguna yang menggunakan sistem ini harus menggunakan PC atau *Laptop* untuk menjalankan sistem prediksi.

3. Analisis Kelayakan Sistem

a. Kelayakan Teknologi

Mengenai kelayakan teknologi, algoritma C4.5 diimplementasikan untuk mengetahui mahasiswa yang dapat lulus tepat waktu, tidak tepat waktu dan *dropout*, diharapkan teknologi ini layak untuk dimanfaatkan.

b. Kelayakan Hukum

Mengenai kelayakan hukum, algoritma C4.5 diimplementasikan untuk prediksi kelulusan mahasiswa menggunakan *software*

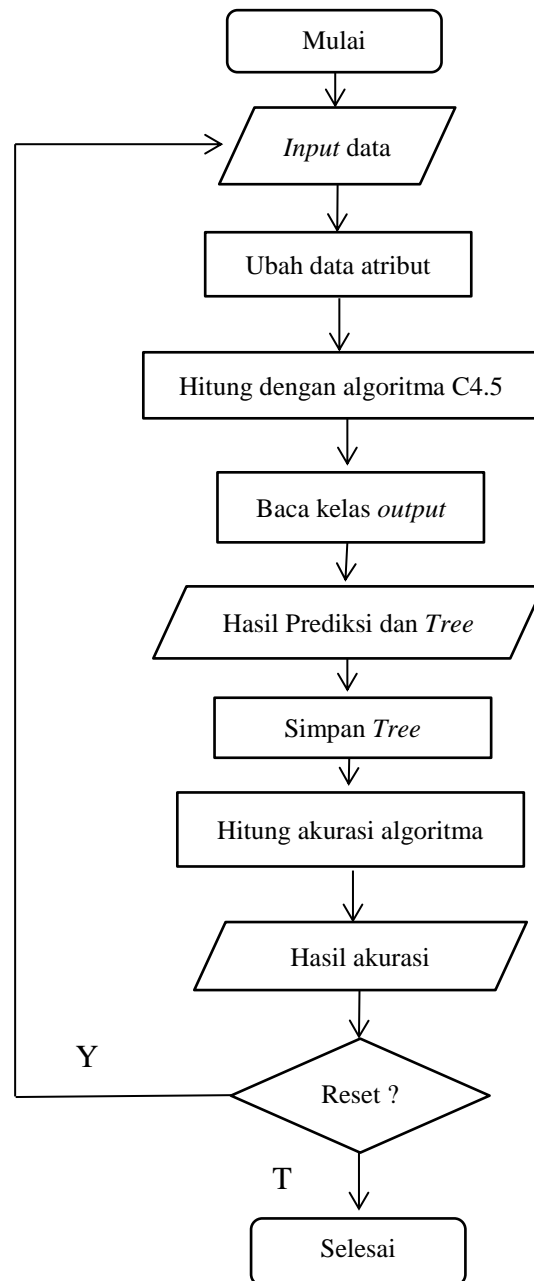
Matlab yang berbayar. *Software* Matlab dapat dibeli dari link <https://www.mathworks.com> baik itu untuk standar, pendidikan, rumah maupun mahasiswa.

c. Kelayakan Operasional

Mengenai kelayakan operasional, sistem yang dibangun akan dirancang lebih sederhana sehingga lebih mudah untuk dipahami saat menggunakan sistem tersebut.

C. Desain Sistem

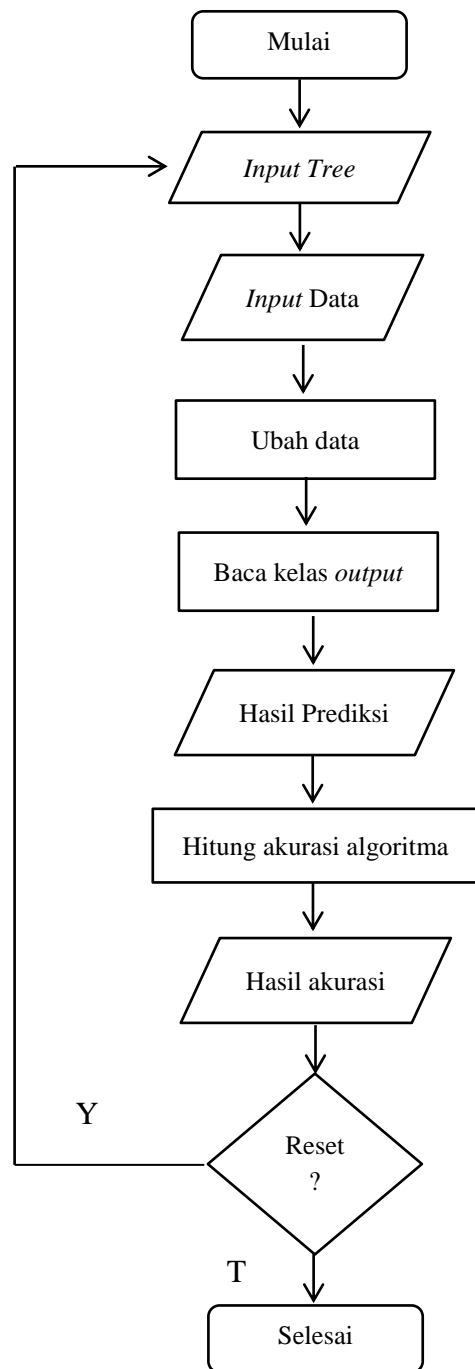
1. Desain Proses



Gambar 7. Flowchart Menu Training

Gambar 7 merupakan *flowchart* pelatihan data yang akan membentuk pohon keputusan untuk prediksi kelulusan mahasiswa.

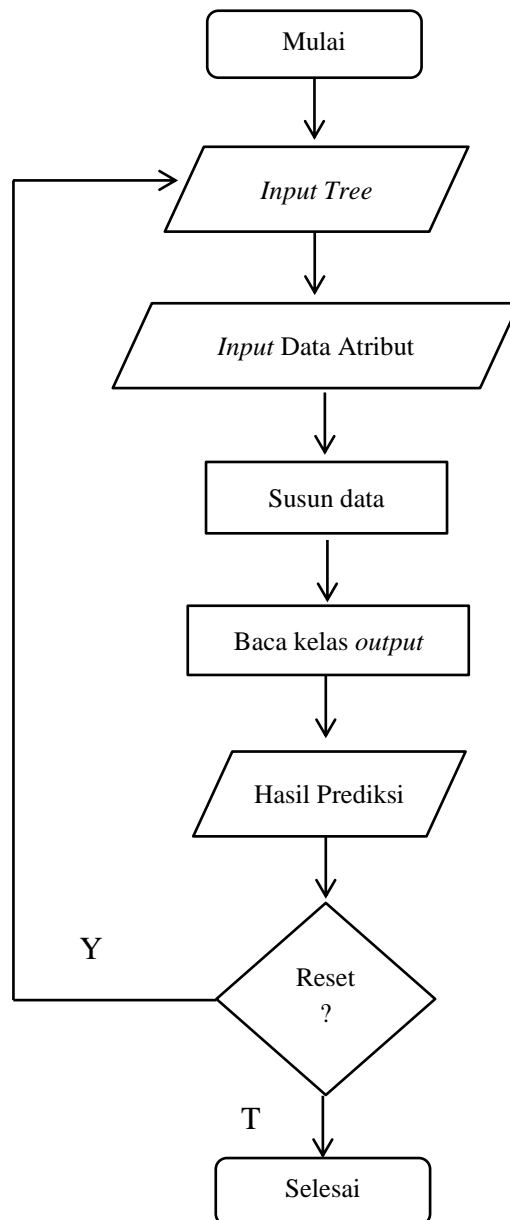
Data *diinput* berupa file excel. Kemudian, data atribut diubah dari kualitatif menjadi kuantitatif khusus untuk data *output* tetap kualitatif yaitu tepat waktu, terlambat dan *drop out*. Setelah data diubah, data dihitung menggunakan algoritma C4.5 lalu sistem akan membaca kelas *output* dan menampilkan hasil dari prediksi dan pohon keputusan. Pohon keputusan hasil pelatihan data kemudian disimpan untuk digunakan dalam pengujian data. Selanjutnya, sistem menghitung akurasi dari algoritma C4.5 yaitu *precision*, *recall*, *f-measure*, dan *accuracy*. Perintah *reset* digunakan untuk mengatur ulang pelatihan data.



Gambar 8. *Flowchart Menu Testing 1*

Gambar 8 merupakan *flowchart* pengujian data berbentuk tabel. Pertama, pohon keputusan hasil pelatihan data yang telah disimpan, diinput. Selanjutnya *input* data yang hendak diuji dan sistem akan

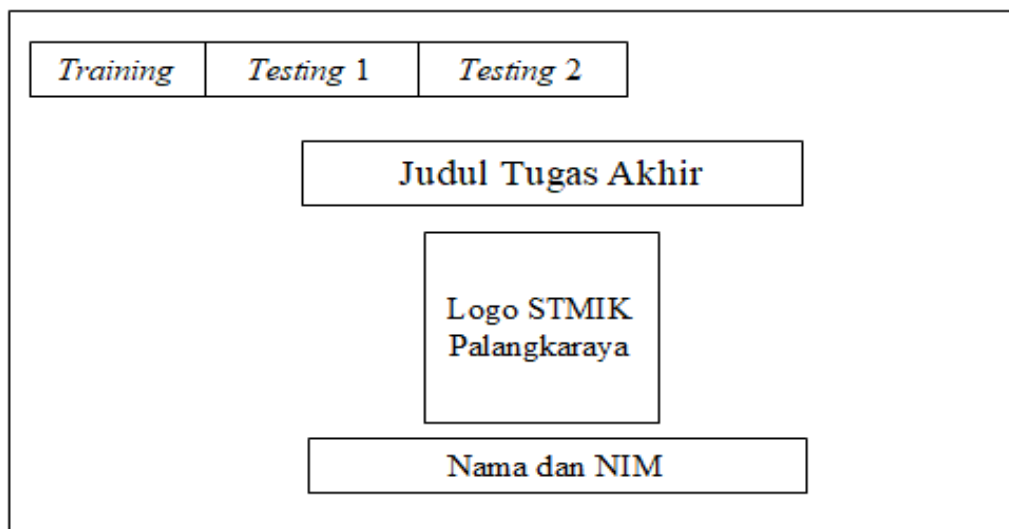
mengubah data atribut menjadi kuantitatif. Setelah data diubah, selanjutnya sistem akan membaca kelas *output* dan menampilkan hasil prediksi dan akurasi algoritma. Perintah *reset* digunakan untuk mengatur ulang pengujian data.



Gambar 9. Flowchart Menu Testing 2

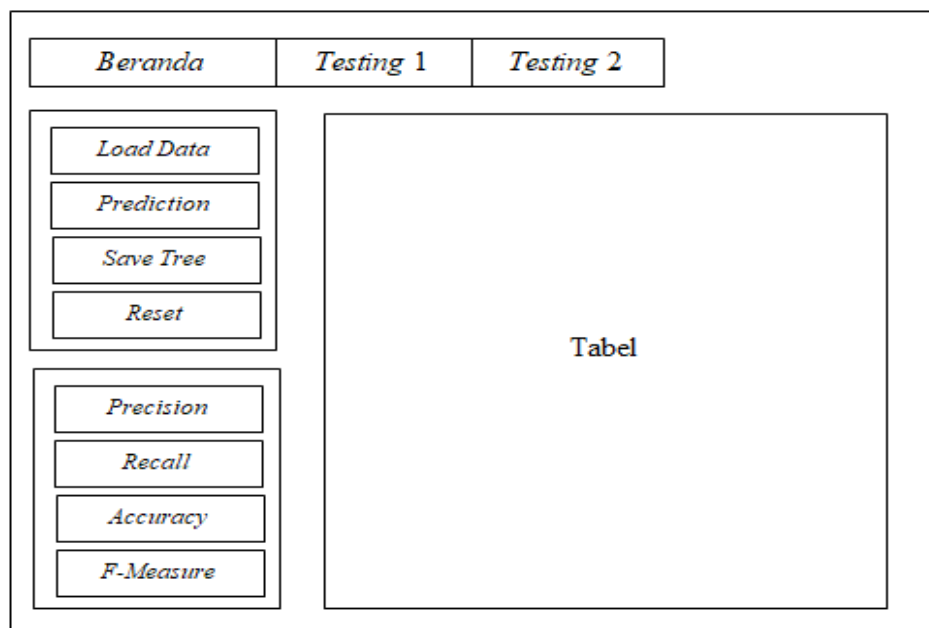
Gambar 9 merupakan *flowchart* pengujian data dalam bentuk *form*. Data-data mahasiswa seperti NIM, nama, jenis kelamin, IPS 1 sampai dengan IPS 6, status pernikahan, status pekerjaan dan jumlah SKS diinput secara langsung tidak melalui *file excel*. Setelah data diinput sistem akan menyusun data atribut menjadi x1 sampai dengan x10 agar dapat dihitung dengan algoritma C4.5. Selanjutnya data diprediksi berdasarkan pohon keputusan yang telah disimpan. Perintah *reset* digunakan untuk mengatur ulang pengujian data.

2. Desain *Interface*/Antarmuka



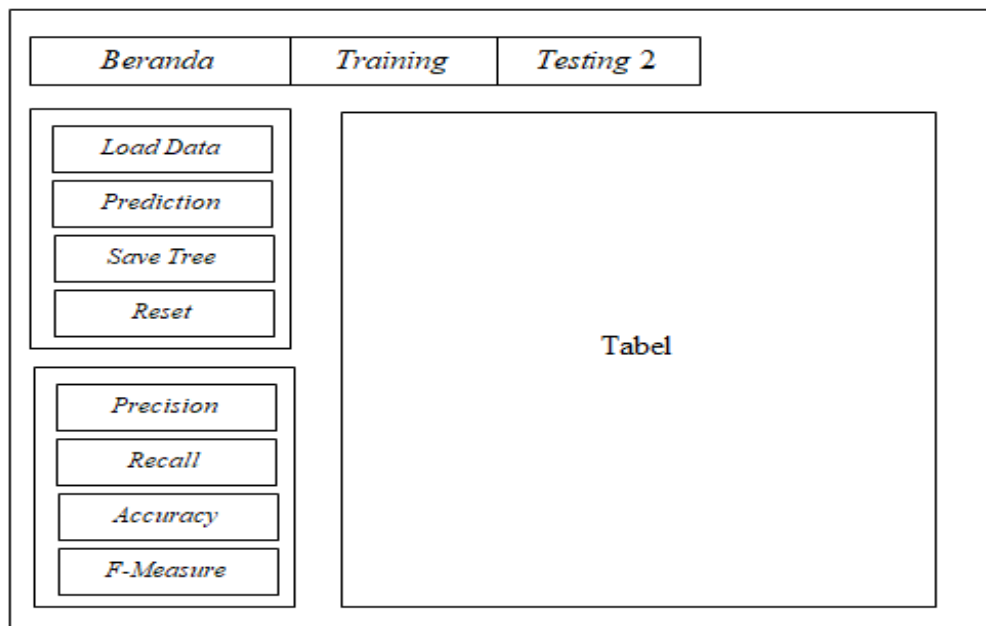
Gambar 10. Desain *Beranda*

Gambar 10 merupakan desain halaman awal pada saat sistem dijalankan.



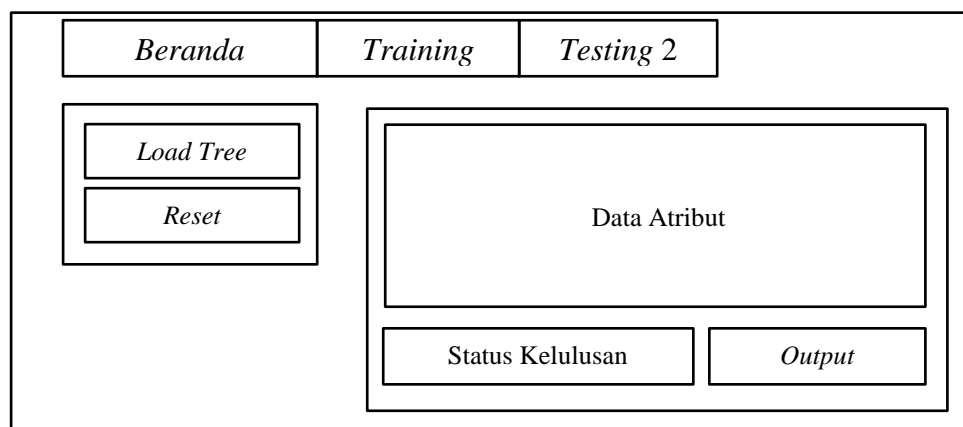
Gambar 11. Desain Menu *Training*

Sebelum melakukan pengujian data terlebih dahulu dilakukan pelatihan data untuk mendapatkan pohon keputusan yang akan digunakan dalam proses pengujian data. Gambar 11 merupakan tampilan menu pelatihan data.



Gambar 12. Desain *Testing 1*

Setelah melakukan pelatihan data, selanjutnya dilakukan pengujian data berdasarkan pohon keputusan yang disimpan dalam proses pelatihan data. Gambar 12 merupakan tampilan menu pengujian data berbentuk tabel.



Gambar 13. Desain *Testing 2*

Gambar 13 merupakan desain pengujian data dalam bentuk *form*. Data atribut yaitu NIM, nama, jenis kelamin, IPS 1, IPS 2,

IPS 3, IPS 4, IPS 5, IPS 6, status pernikahan, status pekerjaan dan jumlah SKS. Tombol prediksi untuk menentukan *output* sistem baik itu tepat waktu, terlambat dan *drop out*.

BAB IV

IMPLEMENTASI DAN PEMBAHASAN

A. Implementasi

1. Uji Coba Sistem dan Program

Uji coba sistem dan program dilakukan dengan pengujian *black box testing* untuk mengetahui apakah sistem telah berjalan sesuai dengan yang diharapkan. Pengujian terbagi menjadi 3 bagian berdasarkan menu yang ada disistem.

a. *Training*

Pada menu *training* pengguna diharuskan untuk menginput *file* excel untuk dihitung menggunakan algoritma C4.5. Pengguna juga diharuskan untuk menyimpan pohon keputusan hasil perhitungan untuk selanjutnya digunakan pada pengujian data.

Tabel 8. Rencana Pengujian Menu *Training*

Kelas Uji	Butir Uji	Jenis Pengujian
Pengujian <i>input file</i> excel dan menampilkannya ditabel yang ada disistem	<i>Load data</i>	<i>Black box</i>
Pengujian proses pelatihan data menggunakan algoritma C4.5	<i>Prediction</i>	<i>Black box</i>
Pengujian penyimpanan pohon keputusan	<i>Save tree</i>	<i>Black box</i>
Mengatur ulang pelatihan data	<i>Reset</i>	<i>Black box</i>

Ketika pengguna menginput *file* excel maka sistem akan melakukan pencarian *file* dan membaca data yang ada di *file* excel tersebut lalu menampilkannya di tabel yang ada di *interface* sistem. Setelah data ditampilkan, selanjutnya sistem akan melakukan

prediksi menggunakan algoritma C4.5 yang menghasilkan *output*, pohon keputusan dan hasil perhitungan akurasi algoritma. Lalu pengguna harus menyimpan pohon keputusan untuk selanjutnya digunakan pada pengujian data. Pengguna dapat menggunakan tombol *reset* untuk mengatur ulang proses pelatihan data.

Tabel 9. Hasil Pengujian Menu *Training*

Data Masukan	Harapan	Pengamatan	Kesimpulan
Menginput <i>file</i> excel	Sistem akan melakukan pencarian file dan membaca data yang ada di <i>file</i> excel lalu menampilkannya di tabel sistem	Pengguna harus mengklik tombol <i>load data</i>	Sesuai
Melakukan pelatihan data	Sistem akan menampilkan hasil prediksi, pohon keputusan dan hasil perhitungan akurasi algoritma.	Pengguna harus mengklik tombol <i>prediction</i>	Sesuai
Menyimpan pohon keputusan	Sistem dapat menyimpan <i>file</i> pohon keputusan	Pengguna harus mengklik tombol <i>save tree</i>	Sesuai
Mengatur ulang proses pelatihan data	Sistem akan mengosongkan tabel, hasil prediksi dan hasil perhitungan akurasi algoritma	Pengguna harus mengklik tombol <i>reset</i>	Sesuai

b. *Testing* 1

Pada menu *testing* 1, pengguna harus melakukan pencarian *file* pohon keputusan agar dapat melakukan prediksi. Selanjutnya pengguna menginput *file* excel untuk diuji.

Tabel 10. Rencana Pengujian Menu *Testing* 1

Kelas Uji	Butir Uji	Jenis Pengujian
Pengujian pencarian <i>file</i> pohon keputusan yang telah disimpan	<i>Load tree</i>	<i>Black box</i>
Pengujian <i>input file</i> excel dan menampilkannya ditabel yang ada disistem	<i>Load data</i>	<i>Black box</i>
Pengujian proses pelatihan data menggunakan algoritma C4.5	<i>Prediction</i>	<i>Black box</i>
Mengatur ulang pengujian data	<i>Reset</i>	<i>Black box</i>

Untuk melakukan pencarian *file* pohon keputusan yang telah disimpan pengguna menggunakan tombol *load tree*. Setelah *file tree* diinput, pengguna menginput *file* excel untuk diuji, melakukan prediksi dan dapat mengatur ulang proses pengujian data.

Tabel 11. Hasil Pengujian Menu *Training* 1

Data Masukan	Harapan	Pengamatan	Kesimpulan
Menginput <i>file tree</i>	Sistem akan melakukan pencarian <i>file tree</i>	Pengguna harus mengklik tombol <i>load tree</i>	Sesuai
Menginput <i>file</i> excel	Sistem akan melakukan pencarian file dan membaca data yang ada di <i>file</i> excel lalu menampilkannya ditabel sistem	Pengguna harus mengklik tombol <i>load data</i>	Sesuai
Melakukan prediksi	Sistem akan menampilkan hasil prediksi dan hasil perhitungan akurasi algoritma.	Pengguna harus mengklik tombol <i>prediction</i>	Sesuai
Mengatur ulang proses pelatihan data	Sistem akan mengosongkan tabel, hasil prediksi dan hasil perhitungan akurasi algoritma	Pengguna harus mengklik tombol <i>reset</i>	Sesuai

c. *Testing 2*

Pada menu *testing 2*, pengujian data dilakukan tidak menggunakan *file excel*. Pada menu ini pengguna *menginput* pohon keputusan yang telah disimpan pada proses pelatihan data. Selanjutnya pengguna akan mengisi data-data atribut yang akan dihitung menggunakan algoritma C4.5 secara langsung.

Tabel 12. Rencana Pengujian Menu *Testing 2*

Kelas Uji	Butir Uji	Jenis Pengujian
Pengujian pencarian <i>file</i> pohon keputusan yang telah disimpan	<i>Load tree</i>	<i>Black box</i>
<i>Menginput</i> data-data atribut	<i>Input data</i>	<i>Black box</i>
Pengujian proses pelatihan data menggunakan algoritma C4.5	<i>Predicton</i>	<i>Black box</i>
Mengatur ulang pengujian data	<i>Reset</i>	<i>Black box</i>

Setelah data-data atribut *diinput*, pengguna melakukan prediksi untuk menghasilkan *output* pengujian data.

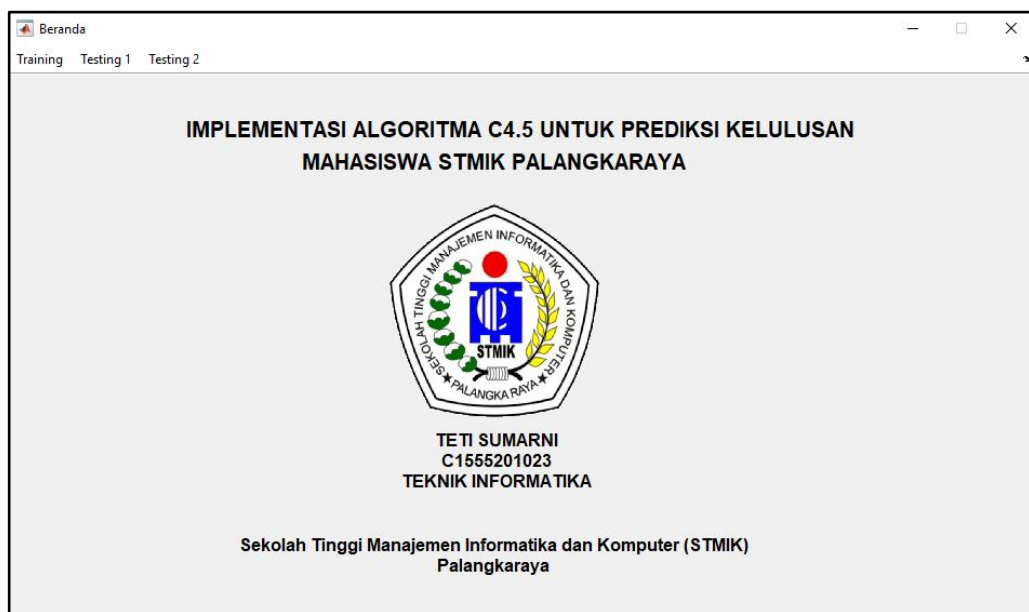
Tabel 13. Hasil Pengujian Menu *Testing 2*

Data Masukan	Harapan	Pengamatan	Kesimpulan
<i>Menginput file tree</i>	Sistem akan melakukan pencarian <i>file tree</i>	Pengguna harus mengklik tombol <i>load tree</i>	Sesuai
<i>Menginput</i> data atribut	Sistem akan membaca data yang telah <i>diinput</i> , jika data tidak terisi dengan benar maka sistem akan menampilkan kotak dialog untuk mngisi data dengan benar	Pengguna harus <i>menginput</i> data secara manual	Sesuai
Melakukan prediksi	Sistem akan menampilkan hasil prediksi	Pengguna harus mengklik tombol <i>prediction</i>	Sesuai, namun sistem menangani jika masukan angka lebih dari 4

Data Masukan	Harapan	Pengamatan	Kesimpulan
Mengatur ulang proses pelatihan data	Sistem akan mengosongkan data <i>inputan</i> dan <i>output</i>	Pengguna harus mengklik tombol <i>reset</i>	Sesuai

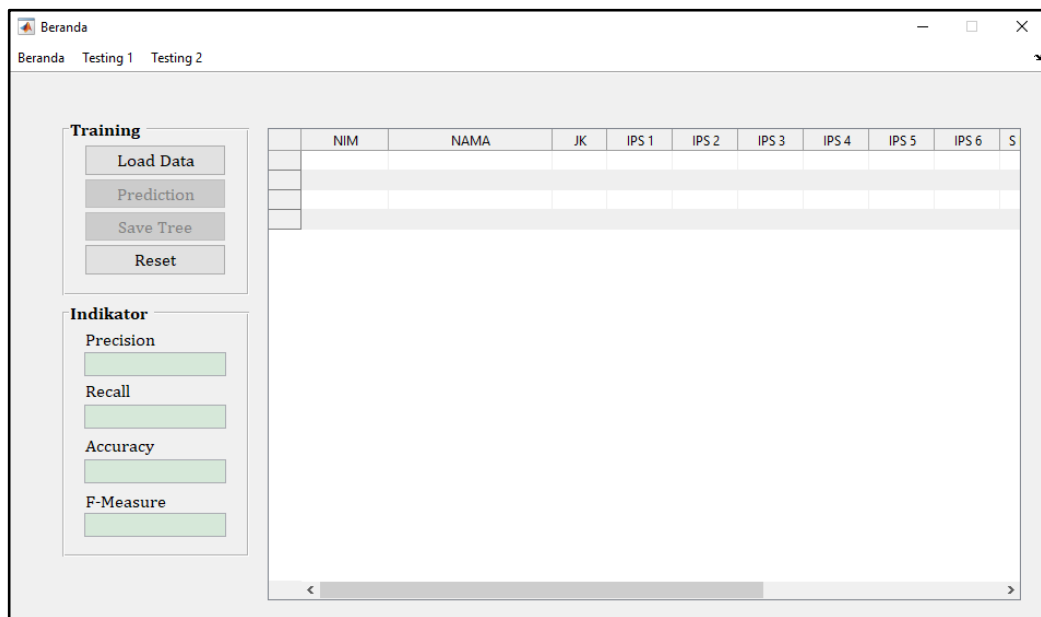
2. Manual Program

Berikut adalah manual penggunaan sistem prediksi kelulusan mahasiswa menggunakan algoritma C4.5.



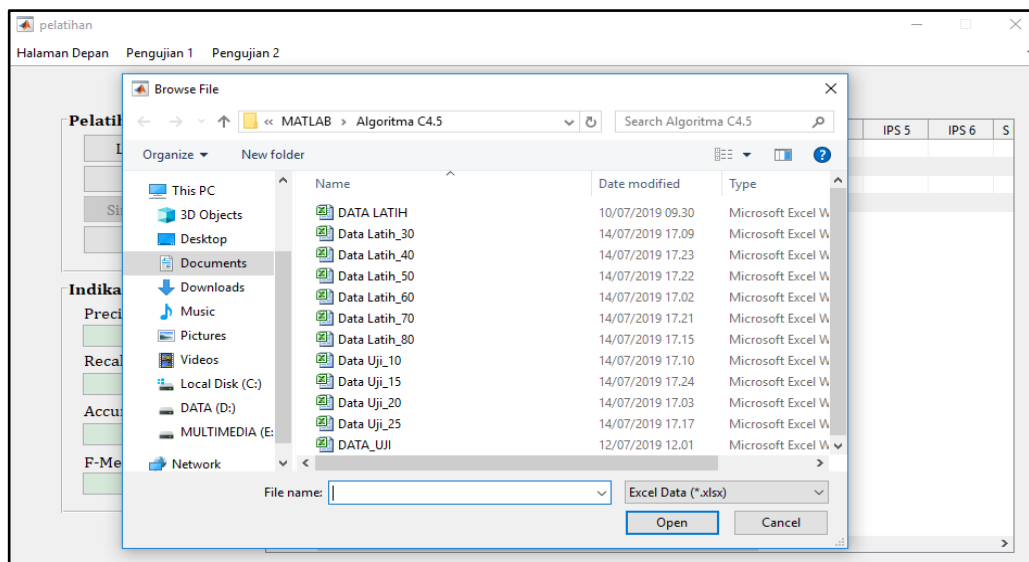
Gambar 14. Tampilan *Beranda*

Gambar 14 merupakan halaman awal sistem prediksi kelulusan mahasiswa pada saat dijalankan. Pada *beranda* terdapat menu *training*, *testing 1* dan *testing 2*.



Gambar 15. Tampilan Menu *Training* 1

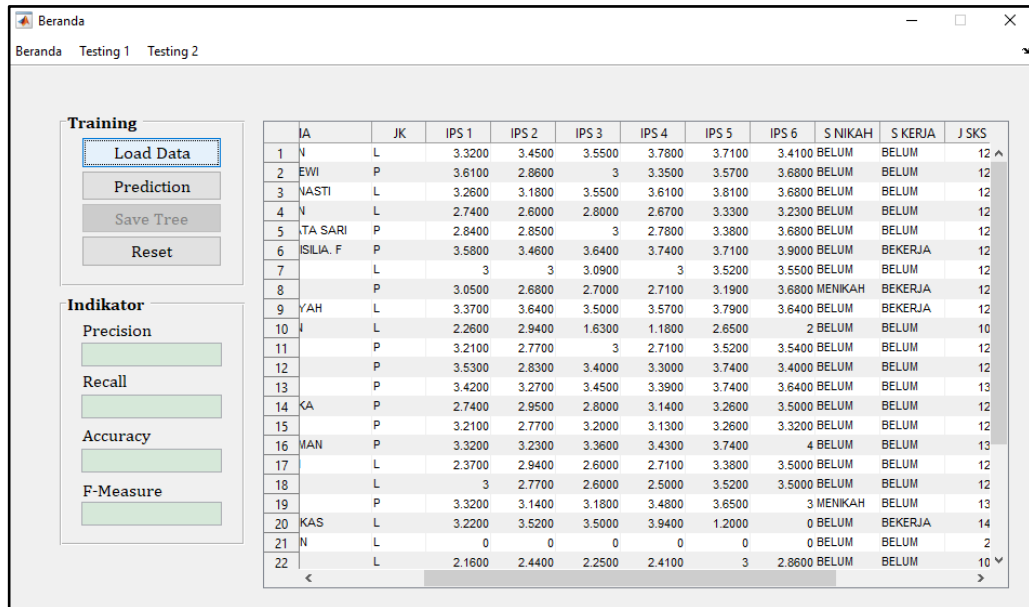
Gambar 15 merupakan menu pelatihan data. Pada menu *training* terdapat tombol *load* data untuk menginput data latih dari *file* excel kedalam tabel.



Gambar 16. Tampilan Pencarian *File*

Gambar 16 merupakan tampilan pencarian *file* excel setelah pengguna mengklik tombol *load* data. Setelah *file* dipilih maka sistem

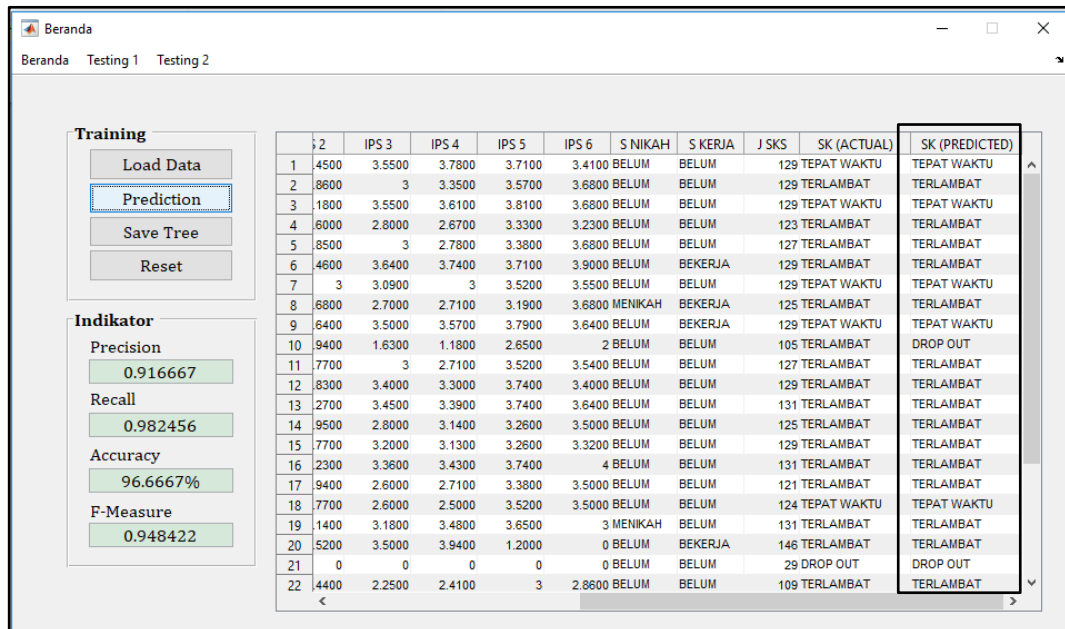
akan membaca jumlah data yang ada di *file* excel lalu menampilkan data tersebut kedalam tabel seperti pada Gambar 17.



The screenshot shows the 'Beranda' application window with the 'Load Data' button highlighted. The table displays data for 22 rows, including columns for ID, Name, Gender, and various IPS scores.

	IA	JK	IPS 1	IPS 2	IPS 3	IPS 4	IPS 5	IPS 6	S NIKAH	S KERJA	J SKS
1	N	L	3.3200	3.4500	3.5500	3.7800	3.7100	3.4100	BELUM	BELUM	12
2	EWI	P	3.6100	2.8600	3	3.3500	3.5700	3.6800	BELUM	BELUM	12
3	NASTI	L	3.2600	3.1800	3.5500	3.6100	3.8100	3.6800	BELUM	BELUM	12
4	N	L	2.7400	2.6000	2.8000	2.6700	3.3300	3.2300	BELUM	BELUM	12
5	TA SARI	P	2.8400	2.8500	3	2.7800	3.3800	3.6800	BELUM	BELUM	12
6	SILIA. F	P	3.5800	3.4600	3.6400	3.7400	3.7100	3.9000	BELUM	BEKERJA	12
7		L	3	3	3.0900	3	3.5200	3.5500	BELUM	BELUM	12
8		P	3.0500	2.6800	2.7000	2.7100	3.1900	3.6800	MENIKAH	BEKERJA	12
9	YAH	L	3.3700	3.6400	3.5000	3.5700	3.7900	3.6400	BELUM	BEKERJA	12
10		L	2.2600	2.9400	1.6300	1.1800	2.6500	2	BELUM	BELUM	10
11		P	3.2100	2.7700	3	2.7100	3.5200	3.5400	BELUM	BELUM	12
12		P	3.5300	2.8300	3.4000	3.3000	3.7400	3.4000	BELUM	BELUM	12
13		P	3.4200	3.2700	3.4500	3.3900	3.7400	3.6400	BELUM	BELUM	13
14	KA	P	2.7400	2.9500	2.8000	3.1400	3.2600	3.5000	BELUM	BELUM	12
15		P	3.2100	2.7700	3.2000	3.1300	3.2600	3.3200	BELUM	BELUM	12
16	MAN	P	3.3200	3.2300	3.3600	3.4300	3.7400	4	BELUM	BELUM	13
17		L	2.3700	2.9400	2.6000	2.7100	3.3800	3.5000	BELUM	BELUM	12
18		L	3	2.7700	2.6000	2.5000	3.5200	3.5000	BELUM	BELUM	12
19		P	3.3200	3.1400	3.1800	3.4800	3.6500	3	MENIKAH	BELUM	13
20	KAS	L	3.2200	3.5200	3.5000	3.9400	1.2000	0	BELUM	BEKERJA	14
21	N	L	0	0	0	0	0	0	BELUM	BELUM	2
22		L	2.1600	2.4400	2.2500	2.4100	3	2.8600	BELUM	BELUM	10

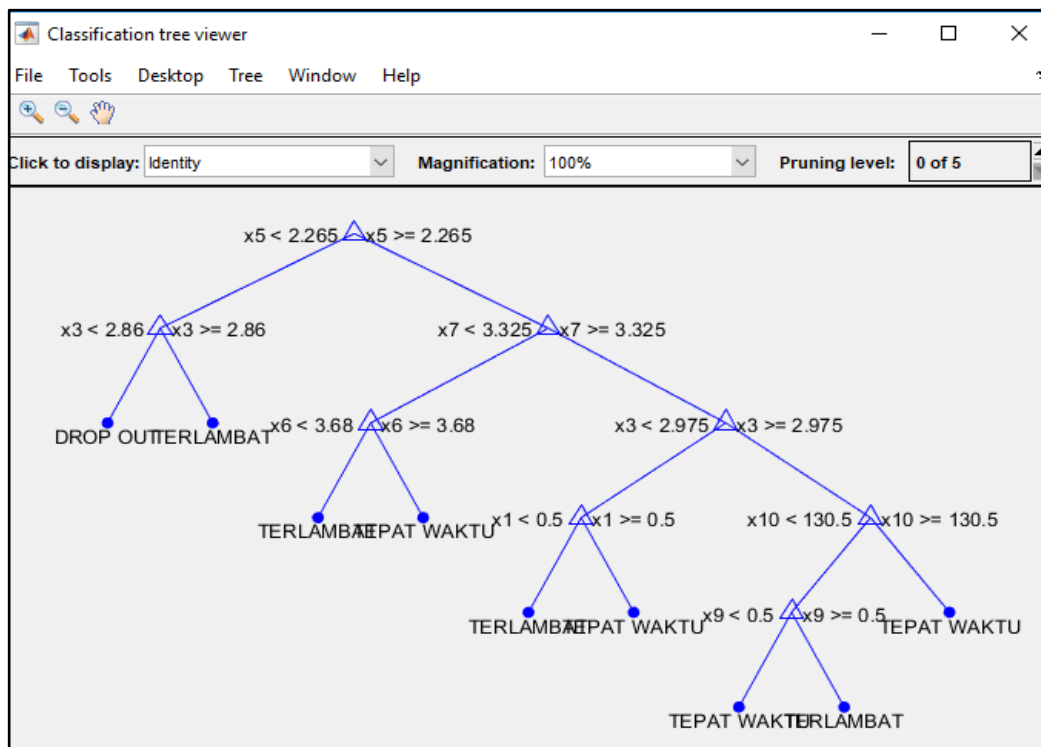
Gambar 17. Tampilan Hasil *Load Data*



The screenshot shows the 'Beranda' application window with the 'Prediction' button highlighted. The table displays predicted results for 22 rows, including columns for ID, IPS scores, and predicted status.

	2	IPS 3	IPS 4	IPS 5	IPS 6	S NIKAH	S KERJA	J SKS	SK (ACTUAL)	SK (PREDICTED)
1	4500	3.5500	3.7800	3.7100	3.4100	BELUM	BELUM	129	TEPAT WAKTU	TEPAT WAKTU
2	8600	3	3.3500	3.5700	3.6800	BELUM	BELUM	129	TERLAMBAT	TERLAMBAT
3	1800	3.5500	3.6100	3.8100	3.6800	BELUM	BELUM	129	TEPAT WAKTU	TEPAT WAKTU
4	6000	2.8000	2.6700	3.3300	3.2300	BELUM	BELUM	123	TERLAMBAT	TERLAMBAT
5	8500	3	2.7800	3.3800	3.6800	BELUM	BELUM	127	TERLAMBAT	TERLAMBAT
6	4600	3.6400	3.7400	3.7100	3.9000	BELUM	BEKERJA	129	TERLAMBAT	TERLAMBAT
7	3	3.0900	3	3.5200	3.5500	BELUM	BELUM	129	TEPAT WAKTU	TEPAT WAKTU
8	6800	2.7000	2.7100	3.1900	3.6800	MENIKAH	BEKERJA	125	TERLAMBAT	TERLAMBAT
9	6400	3.5000	3.5700	3.7900	3.6400	BELUM	BEKERJA	129	TEPAT WAKTU	TEPAT WAKTU
10	9400	1.6300	1.1800	2.6500	2	BELUM	BELUM	105	TERLAMBAT	DROP OUT
11	7700	3	2.7100	3.5200	3.5400	BELUM	BELUM	127	TERLAMBAT	TERLAMBAT
12	8300	3.4000	3.3000	3.7400	3.4000	BELUM	BELUM	129	TERLAMBAT	TERLAMBAT
13	2700	3.4500	3.3900	3.7400	3.6400	BELUM	BELUM	131	TERLAMBAT	TERLAMBAT
14	9500	2.8000	3.1400	3.2600	3.5000	BELUM	BELUM	125	TERLAMBAT	TERLAMBAT
15	7700	3.2000	3.1300	3.2600	3.3200	BELUM	BELUM	129	TERLAMBAT	TERLAMBAT
16	2300	3.3600	3.4300	3.7400	4	BELUM	BELUM	131	TERLAMBAT	TERLAMBAT
17	9400	2.6000	2.7100	3.3800	3.5000	BELUM	BELUM	121	TERLAMBAT	TERLAMBAT
18	7700	2.6000	2.5000	3.5200	3.5000	BELUM	BELUM	124	TEPAT WAKTU	TEPAT WAKTU
19	1400	3.1800	3.4800	3.6500	3	MENIKAH	BELUM	131	TERLAMBAT	TERLAMBAT
20	5200	3.5000	3.9400	1.2000	0	BELUM	BEKERJA	146	TERLAMBAT	TERLAMBAT
21	0	0	0	0	0	BELUM	BELUM	29	DROP OUT	DROP OUT
22	4400	2.2500	2.4100	3	2.8600	BELUM	BELUM	109	TERLAMBAT	TERLAMBAT

Gambar 18. Tampilan Hasil Prediksi



Gambar 19. Tampilan Pohon Keputusan

Training

Load Data

Prediction

Save Tree

Reset

Indikator

Precision: 0.916667

Recall: 0.982456

Accuracy: 96.6667%

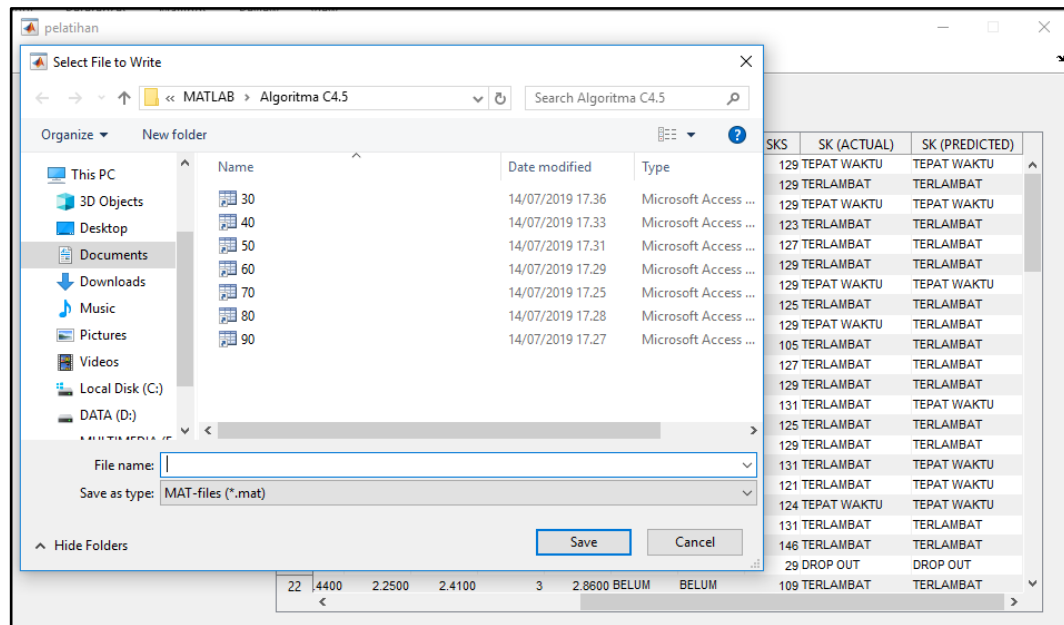
F-Measure: 0.948422

	2	IPS 3	IPS 4	IPS 5	IPS 6	S NIKAH	S KERJA	J SKS	SK (ACTUAL)	SK (PREDICTED)
1	4500	3.5500	3.7800	3.7100	3.4100	BELUM	BELUM	129	TEPAT WAKTU	TEPAT WAKTU
2	8600	3	3.3500	3.5700	3.6800	BELUM	BELUM	129	TERLAMBAT	TERLAMBAT
3	1800	3.5500	3.6100	3.8100	3.6800	BELUM	BELUM	129	TEPAT WAKTU	TEPAT WAKTU
4	6000	2.8000	2.6700	3.3300	3.2300	BELUM	BELUM	123	TERLAMBAT	TERLAMBAT
5	8500	3	2.7800	3.3800	3.6800	BELUM	BELUM	127	TERLAMBAT	TERLAMBAT
6	4600	3.6400	3.7400	3.7100	3.9000	BELUM	BEKERJA	129	TERLAMBAT	TERLAMBAT
7	3	3.0900	3	3.5200	3.5500	BELUM	BELUM	129	TEPAT WAKTU	TEPAT WAKTU
8	8800	2.7000	2.7100	3.1900	3.6800	MENIKAH	BEKERJA	125	TERLAMBAT	TERLAMBAT
9	6400	3.5000	3.5700	3.7900	3.6400	BELUM	BEKERJA	129	TEPAT WAKTU	TEPAT WAKTU
10	9400	1.6300	1.1800	2.6500	2	BELUM	BELUM	105	TERLAMBAT	DROP OUT
11	7700	3	2.7100	3.5200	3.5400	BELUM	BELUM	127	TERLAMBAT	TERLAMBAT
12	8300	3.4000	3.3000	3.7400	3.4000	BELUM	BELUM	129	TERLAMBAT	TERLAMBAT
13	2700	3.4500	3.3900	3.7400	3.6400	BELUM	BELUM	131	TERLAMBAT	TERLAMBAT
14	9500	2.8000	3.1400	3.2600	3.5000	BELUM	BELUM	125	TERLAMBAT	TERLAMBAT
15	7700	3.2000	3.1300	3.2600	3.3200	BELUM	BELUM	129	TERLAMBAT	TERLAMBAT
16	2300	3.3600	3.4300	3.7400	4	BELUM	BELUM	131	TERLAMBAT	TERLAMBAT
17	9400	2.6000	2.7100	3.3800	3.5000	BELUM	BELUM	121	TERLAMBAT	TERLAMBAT
18	7700	2.6000	2.5000	3.5200	3.5000	BELUM	BELUM	124	TEPAT WAKTU	TEPAT WAKTU
19	1400	3.1800	3.4800	3.6500	3	MENIKAH	BELUM	131	TERLAMBAT	TERLAMBAT
20	5200	3.5000	3.9400	1.2000	0	BELUM	BEKERJA	146	TERLAMBAT	TERLAMBAT
21	0	0	0	0	0	BELUM	BELUM	29	DROP OUT	DROP OUT
22	4400	2.2500	2.4100	3	2.8600	BELUM	BELUM	109	TERLAMBAT	TERLAMBAT

Gambar 20. Hasil Akurasi Algoritma C4.5

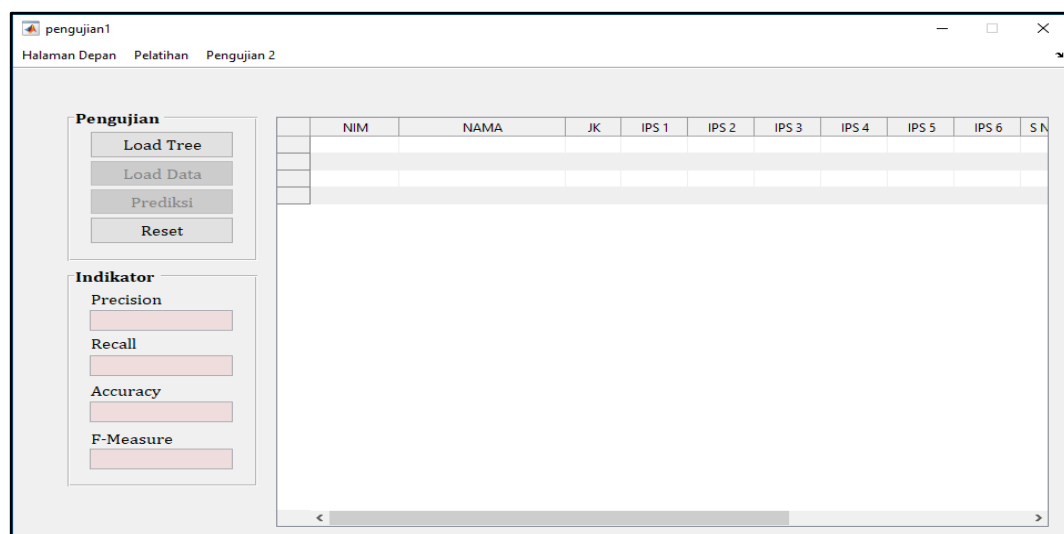
Tombol *prediction* pada menu pelatihan berfungsi untuk menampilkan hasil prediksi seperti pada Gambar 18, pohon keputusan

seperti pada Gambar 19 dan akurasi algoritma (*precision*, *recall*, *accuracy* dan *f-measure*) seperti pada Gambar 20.



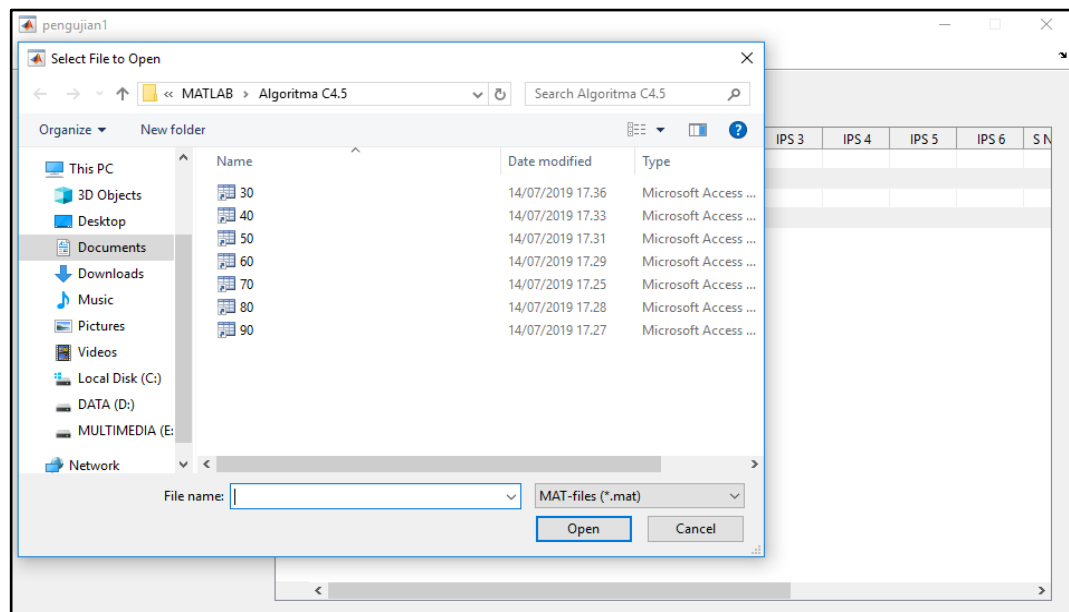
Gambar 21. Tampilan Simpan *Tree*

Tombol *save tree* untuk menyimpan pohon keputusan seperti pada Gambar 21. Tombol *reset* untuk mengatur ulang pelatihan data sehingga sistem akan kembali seperti pada Gambar 15.



Gambar 22. Tampilan Menu *Testing* 1

Gambar 22 merupakan pengujian data dalam bentuk tabel. Tombol *load tree* untuk menginput pohon keputusan yang telah disimpan seperti pada Gambar 23.



Gambar 23. Tampilan *Load Tree*

Gambar 24. Tampilan Menu *Testing 2*

The screenshot shows a web application window titled "pengujian2" with a navigation bar containing "Halaman Depan", "Pelatihan", and "Pengujian 1". The main content area is divided into two sections: "Pengujian" and "Form".

The "Pengujian" section contains two buttons: "Load Tree" and "Reset".

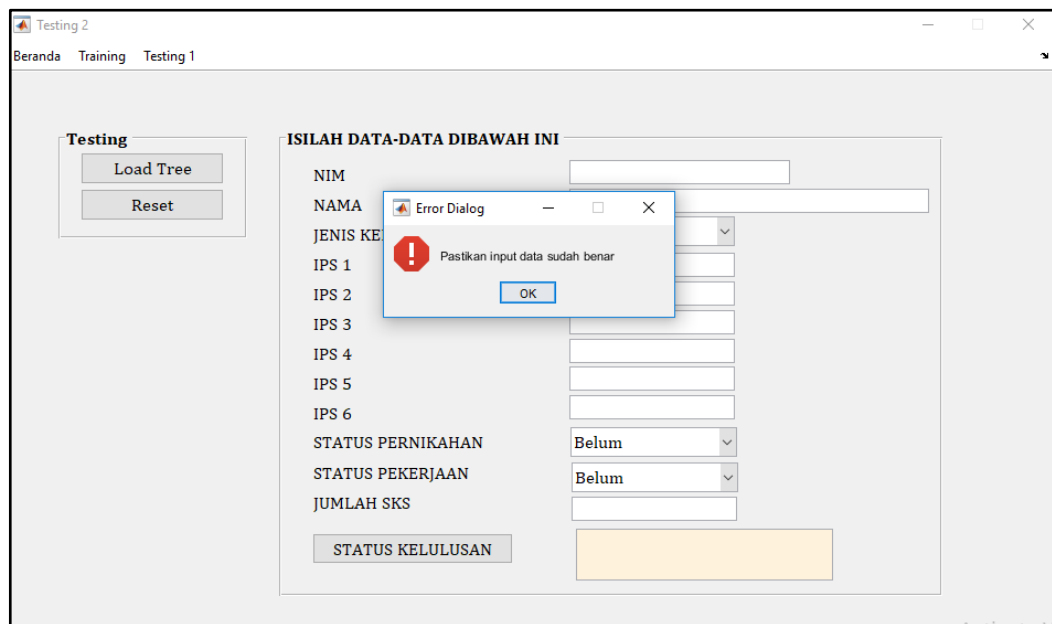
The "Form" section contains the following fields and values:

Field	Value
NIM	C1455201021
NAMA	PEBRIANSYAH
JENIS KELAMIN	Laki-Laki
IPS 1	2,95
IPS 2	2,86
IPS 3	2,70
IPS 4	3,27
IPS 5	3,09
IPS 6	3,57
S NIKAH	Belum
S KERJA	Belum
J SKS	127

At the bottom of the form, there are two buttons: "SK" (blue) and "TEPAT WAKTU" (yellow).

Gambar 25. Tampilan Hasil Prediksi

Gambar 24 merupakan pengujian data dalam bentuk *form*. Data-data atribut yaitu NIM, nama, jenis kelamin, IPS 1, IPS 2, IPS 3, IPS 4, IPS 5, IPS 6, status pernikahan, status pekerjaan dan jumlah SKS diinput secara langsung tidak melalui *file excel*. Selanjutnya data diprediksi berdasarkan pohon keputusan yang telah disimpan untuk menghasilkan *output* status kelulusan tepat waktu, terlambat atau *drop out* seperti pada Gambar 25.



Gambar 26. Tampilan Kotak Dialog

Gambar 26 akan muncul jika data mahasiswa yang *diinput* kosong.

B. Pembahasan

1. Pembahasan Listing Program

Berikut ini merupakan pembahasan listing program yang digunakan dalam membangun sistem prediksi kelulusan mahasiswa.

a. *Load Data*

```
% menampilkan menu browse file
[nama_file,nama_path] = uigetfile({'*.xlsx','Excel Data (*.xlsx)'};
    '*.','All file type (*.*)'},'Browse File');

% jika ada file yang dipilih maka akan menjalankan perintah
% di bawahnya
if ~isequal(nama_file,0)
    % membaca data yang ada di file excel
    [~, ~, alldata] = xlsread(fullfile(nama_path,nama_file),1);
    % membaca jumlah data
    numdata = size(alldata,1);
    % menginisialisasi variabel data_mahasiswa
    data_mahasiswa = cell(numdata-1,13);
    for n = 2:numdata
        data_mahasiswa{n-1,1} = alldata{n,1};
```

```

        data_mahasiswa{n-1,2} = alldata{n,2};
        data_mahasiswa{n-1,3} = alldata{n,3};
        data_mahasiswa{n-1,4} = alldata{n,4};
        data_mahasiswa{n-1,5} = alldata{n,5};
        data_mahasiswa{n-1,6} = alldata{n,6};
        data_mahasiswa{n-1,7} = alldata{n,7};
        data_mahasiswa{n-1,8} = alldata{n,8};
        data_mahasiswa{n-1,9} = alldata{n,9};
        data_mahasiswa{n-1,10} = alldata{n,10};
        data_mahasiswa{n-1,11} = alldata{n,11};
        data_mahasiswa{n-1,12} = alldata{n,12};
        data_mahasiswa{n-1,13} = alldata{n,13};
    end

    % menampilkan data pada tabel

    set(handles.uitable1, 'Data', data_mahasiswa, 'RowName', 1:size(data_mahasiswa,1))

    % menyimpan variabel data_mahasiswa pada lokasi
    handles agar dapat
    % dipanggil oleh pushbutton yang lain
    handles.data_mahasiswa = data_mahasiswa;
    guidata(hObject, handles);
else
    % jika tidak ada file yang dipilih maka akan kembali
    return
end

```

Sebelum data dihitung menggunakan algoritma C4.5, data diinput berupa *file* excel. Sistem akan membaca *file* excel dan jumlah data yang ada didalam *file* excel. Sistem akan menginisialisasikan variabel data mahasiswa dan menampilkannya pada tabel yang ada disistem.

b. Ubah Data

```

% pre-processing data
for n = 1:numdata
    if isequal(data_mahasiswa{n,3}, 'L');
        atribut(n,1) = 1;
    elseif isequal(data_mahasiswa{n,3}, 'P');
        atribut(n,1) = 0;
    end

    atribut(n,2) = data_mahasiswa{n,4};
    atribut(n,3) = data_mahasiswa{n,5};
    atribut(n,4) = data_mahasiswa{n,6};
    atribut(n,5) = data_mahasiswa{n,7};
    atribut(n,6) = data_mahasiswa{n,8};
    atribut(n,7) = data_mahasiswa{n,9};

```

```

if isequal(data_mahasiswa{n,10},'BELUM');
    atribut(n,8) = 0;
elseif isequal(data_mahasiswa{n,10},'MENIKAH');
    atribut(n,8) = 1;
end

if isequal(data_mahasiswa{n,11},'BELUM');
    atribut(n,9) = 0;
elseif isequal(data_mahasiswa{n,11},'BEKERJA');
    atribut(n,9) = 1;
end

atribut(n,10) = data_mahasiswa{n,12};

target{n,1} = data_mahasiswa{n,13};
end

```

c. Prediksi

Setelah data diubah, selanjutnya sistem akan menghitung data menggunakan algoritma C4.5 dengan perintah seperti dibawah ini.

```

% prediksi menggunakan algoritma C4.5
t = classregtree(atribut,target);

```

Setelah proses perhitungan selesai, sistem akan membaca kelas *output*, menampilkan pohon keputusan dan menampilkan hasil prediksi.

```

% membaca kelas keluaran hasil prediksi
isLabels = eval(t,atribut);
% menampilkan pohon keputusan
view(t)
% menampilkan data_mahasiswa beserta hasil prediksi pada
tabel
data_mahasiswa = [data_mahasiswa,isLabels];
set(handles.uitable1,'Data',data_mahasiswa,'RowName',1:size(data_mahasiswa,1))

```

d. Simpan Tree

Selanjutnya sistem akan menyimpan pohon keputusan menggunakan source code dibawah ini.

```

% menampilkan menu save file
[nama_file,~] = uiputfile({'*.mat'});

```

e. *Load Tree*

Untuk melakukan pengujian data terlebih dahulu input pohon keputusan yang telah disimpan menggunakan source code dibawah ini

```
load(fullfile(nama_path,nama_file))
```

f. *Hitung Akurasi Algoritma*

```
% confusion matrix
confMat = confusionmat(target,isLabels);

%%% precision
for i = 1:size(confMat,1)
    precision(i) = confMat(i,i)/sum(confMat(:,i));
end
Precision = sum(precision)/size(confMat,1);
% menampilkan precision pada edit text
set(handles.edit1,'String',Precision)

% recall
for i =1:size(confMat,1)
    recall(i)=confMat(i,i)/sum(confMat(i,:));
end
recall(isnan(recall))=[];
Recall = sum(recall)/size(confMat,1);
% menampilkan recall pada edit text
set(handles.edit2,'String',Recall)

%%% Accuracy
Accuracy =
sum(strcmp(isLabels,target))/numel(target)*100;
% menampilkan Accuracy pada edit text
set(handles.edit3,'String',[num2str(Accuracy),'%'])

%%% F-score
F_score = 2*Recall*Precision/(Precision+Recall);
%%F_score=2*1/((1/Precision)+(1/Recall));
% menampilkan F-score pada edit text
```

g. *Reset*

Setelah selesai melakukan prediksi dan perhitungan akurasi algoritma, sistem dapat mengatur ulang sistem menggunakan source code dibawah ini.

```

% mereset button2
set(handles.pushbutton2,'Enable','off')
set(handles.pushbutton3,'Enable','off')
set(handles.uitable1,'Data',[])
set(handles.edit1,'String',[])
set(handles.edit2,'String',[])
set(handles.edit3,'String',[])
set(handles.edit4,'String',[])

```

h. Baca Data yang ada pada *Form*

```

% membaca data2 yg ada pada form
val1 = get(handles.popupmenu1,'Value');
switch val1
    case 1
        x1 = 1;
    case 2
        x1 = 0;
end

x2 = str2double(get(handles.edit3,'String'));
x3 = str2double(get(handles.edit4,'String'));
x4 = str2double(get(handles.edit5,'String'));
x5 = str2double(get(handles.edit6,'String'));
x6 = str2double(get(handles.edit7,'String'));
x7 = str2double(get(handles.edit8,'String'));

val2 = get(handles.popupmenu2,'Value');
switch val2
    case 1
        x8 = 0;
    case 2
        x8 = 1;
end

val3 = get(handles.popupmenu3,'Value');
switch val3
    case 1
        x9 = 0;
    case 2
        x9 = 1;
end

x10 = str2double(get(handles.edit9,'String'));
% menggabungkan seluruh input
atribut = [x1,x2,x3,x4,x5,x6,x7,x8,x9,x10];

```

Setelah data diinput secara manual, sistem akan membaca inputan data dan menyusunnya.

i. Kotak Dialog

Jika terdapat *inputan* data yang kosong maka sistem akan menampilkan kotak dialog menggunakan source code dibawah ini.

```
errordlg('Pastikan input data sudah benar');
```

2. Pembahasan Hasil Penelitian

Pada pembahasan hasil penelitian, peneliti menggunakan data simulasi untuk pelatihan data dan pengujian.

Tabel 14. Data Latih

NIM	NAMA	JK	IPS_1	IPS_2	IPS_3	IPS_4	IPS_5	IPS_6	S_NIKAH	S_KERJA	J_SKS	SK
C1255201001	ROBY SETIAWAN	L	3,32	3,45	3,55	3,78	3,71	3,41	BELUM	BELUM	129	TEPAT WAKTU
C1255201003	RADEN RIZKY PINASTI	L	3,26	3,18	3,55	3,61	3,81	3,68	BELUM	BELUM	129	TEPAT WAKTU
C1255201005	EDY KURNIAWAN	L	2,74	2,60	2,80	2,67	3,33	3,23	BELUM	BELUM	123	TERLAMBAT
C1255201006	YULIANA PERMATA SARI	P	2,84	2,85	3	2,78	3,38	3,68	BELUM	BELUM	127	TERLAMBAT
C1255201023	ERWIN SETIAWAN	L	0	0	0	0	0	0	BELUM	BELUM	29	DROP OUT
C1255201053	BENY BAGUNA	L	0	0	0	0	0	0	MENIKAH	BELUM	19	DROP OUT
C1355201004	JEKIE PRANATHA	L	2,76	2,71	2,91	3,36	3,54	3,44	BELUM	BELUM	128	TEPAT WAKTU
C1355201005	ROBINSON	L	2,95	3,05	3,08	3,36	3,82	3,5	BELUM	BELUM	130	TEPAT WAKTU
C1355201008	VINO AGUSTIANTO.P	L	2,57	2,38	2,44	2,15	0	0	BELUM	BELUM	80	DROP OUT
C1355201033	SANUSI FADLI	P	2,67	2,57	2,83	3,09	2,96	2,75	BELUM	BELUM	132	TERLAMBAT
C1355201039	HEPPY AYU. A	P	3,05	2,74	2,82	3,25	3,55	3,39	BELUM	BELUM	126	TERLAMBAT
C1355201044	RIYADI	L	0,29	0	0	0	0	0	BELUM	BELUM	38	DROP OUT
C1455201008	KLARA DWI LIGWINA	P	2,86	2,86	3,39	3,75	3,54	3,64	BELUM	BEKERJA	133	TEPAT WAKTU
C1455201036	KHUSNUL DELIMAWATI	P	3,38	3,13	3	3,3	3,32	3,7	MENIKAH	BELUM	126	TERLAMBAT
C1455201041	PENIE	P	0,38	0	0	2	0	0	BELUM	BELUM	45	DROP OUT

Tabel 14 merupakan data yang akan digunakan dalam proses pelatihan untuk membentuk pohon keputusan yang akan digunakan pada pengujian.

Tabel 15. Tranformasi Data

JK	IPS1	IPS2	IPS3	IPS4	IPS5	IPS6	S NIKAH	S KERJA	J SKS	SK
1	3,32	3,45	3,55	3,78	3,71	3,41	0	0	129	TEPAT WAKTU
1	3,26	3,18	3,55	3,61	3,81	3,68	0	0	129	TEPAT WAKTU
1	2,74	2,60	2,80	2,67	3,33	3,23	0	0	123	TERLAMBAT
0	2,84	2,85	3	2,78	3,38	3,68	0	0	127	TERLAMBAT
1	0	0	0	0	0	0	0	0	29	DROP OUT
1	0	0	0	0	0	0	1	0	19	DROP OUT
1	2,76	2,71	2,91	3,36	3,54	3,44	0	0	128	TEPAT WAKTU
1	2,95	3,05	3,08	3,36	3,82	3,5	0	0	130	TEPAT WAKTU
1	2,57	2,38	2,44	2,15	0	0	0	0	80	DROP OUT
0	2,67	2,57	2,83	3,09	2,96	2,75	0	0	132	TERLAMBAT
0	3,05	2,74	2,82	3,25	3,55	3,39	0	0	126	TERLAMBAT
1	0,29	0	0	0	0	0	0	0	38	DROP OUT
0	2,86	2,86	3,39	3,75	3,54	3,64	0	1	133	TEPAT WAKTU
0	3,38	3,13	3	3,3	3,32	3,7	1	0	126	TERLAMBAT
0	0,38	0	0	2	0	0	0	0	45	DROP OUT

Tabel 15 merupakan tranformasi data atribut yang kualitatif menjadi kuantitatif yaitu atribut jenis kelamin, status pernikahan dan status pekerjaan. Untuk status kelulusan tetap kualitatif yaitu tepat waktu, terlambat dan *drop out*. Setelah data diubah, selanjutnya data dihitung menggunakan algoritma C4.5 untuk membentuk pohon keputusan. Sebelum proses perhitungan, peneliti mencari nilai *threshold* atau nilai batas parameter untuk masing-masing atribut. Untuk mencari nilai *threshold* yang tepat untuk digunakan sebagai nilai batas digunakan Persamaan :

$$Info(T) = - \sum_{i=1}^k \left(\left(\frac{frequency(C_i, T)}{|T|} \right) \right) \cdot \log_2 \left(\left(\frac{frequency(C_i, T)}{|T|} \right) \right) \quad (3)$$

Frequency(C_i, T) jumlah sampel di himpunan T.

$$\begin{aligned}
 Info(T) &= \left(-\left(\frac{5}{|S|}\right)\right) \cdot \log_2 \left(\left(\frac{5}{|S|}\right)\right) + \left(-\left(\frac{5}{|S|}\right)\right) \cdot \log_2 \left(\left(\frac{5}{|S|}\right)\right) + \left(-\left(\frac{5}{|S|}\right)\right) \cdot \\
 &\log_2 \left(\left(\frac{5}{|S|}\right)\right) \\
 &= 1,584962501
 \end{aligned}$$

Selanjutnya, nilai setiap diurutkan dari nilai terendah ke nilai tertinggi dan dicari nilai *threshold* = nilai atribut pertama- nilai atribut kedua lalu dibagi 2.

Tabel 16. Nilai *Threshold* Atribut Jenis Kelamin

JK	SK	<i>Threshold</i>
0	TERLAMBAT	0,0
0	TERLAMBAT	0,0
0	TERLAMBAT	0,0
0	TEPAT WAKTU	0,0
0	TERLAMBAT	0,0
0	DROP OUT	0,5
1	TEPAT WAKTU	1,0
1	TEPAT WAKTU	1,0
1	TERLAMBAT	1,0
1	DROP OUT	1,0
1	DROP OUT	1,0
1	TEPAT WAKTU	1,0
1	TEPAT WAKTU	1,0
1	DROP OUT	1,0
1	DROP OUT	

Tabel 17. Nilai *Threshold* Atribut IPS 1

IPS_1	SK	<i>Theshold</i>
0	DROP OUT	0,00
0	DROP OUT	0,15
0,29	DROP OUT	0,34
0,38	DROP OUT	1,48
2,57	DROP OUT	2,62
2,67	TERLAMBAT	2,71
2,74	TERLAMBAT	2,75
2,76	TEPAT WAKTU	2,80
2,84	TERLAMBAT	2,85

IPS_1	SK	<i>Theshold</i>
2,86	TEPAT WAKTU	2,91
2,95	TEPAT WAKTU	3,00
3,05	TERLAMBAT	3,16
3,26	TEPAT WAKTU	3,29
3,32	TEPAT WAKTU	3,35
3,38	TERLAMBAT	

Tabel 18. Nilai *Threshold* Atribut IPS 2

IPS_2	SK	<i>Theshold</i>
0	DROP OUT	0,00
0	DROP OUT	0,00
0	DROP OUT	0,00
0	DROP OUT	1,19
2,38	DROP OUT	2,48
2,57	TERLAMBAT	2,59
2,60	TERLAMBAT	2,66
2,71	TEPAT WAKTU	2,73
2,74	TERLAMBAT	2,80
2,85	TERLAMBAT	2,86
2,86	TEPAT WAKTU	2,96
3,05	TEPAT WAKTU	3,09
3,13	TERLAMBAT	3,16
3,18	TEPAT WAKTU	3,32
3,45	TEPAT WAKTU	

Tabel 19. Nilai *Threshold* Atribut IPS 3

IPS_3	SK	<i>Theshold</i>
0	DROP OUT	0,00
0	DROP OUT	0,00
0	DROP OUT	0,00
0	DROP OUT	1,22
2,44	DROP OUT	2,62
2,80	TERLAMBAT	2,81
2,82	TERLAMBAT	2,83
2,83	TERLAMBAT	2,87
2,91	TEPAT WAKTU	2,96
3	TERLAMBAT	3,00
3	TERLAMBAT	3,04
3,08	TEPAT WAKTU	3,24
3,39	TEPAT WAKTU	3,47

IPS_3	SK	<i>Threshold</i>
3,55	TEPAT WAKTU	3,55
3,55	TEPAT WAKTU	

Tabel 20. Nilai *Threshold* Atribut IPS 4

IPS_4	SK	<i>Threshold</i>
0	DROP OUT	0,00
0	DROP OUT	0,00
0	DROP OUT	1,00
2	DROP OUT	2,08
2,15	DROP OUT	2,41
2,67	TERLAMBAT	2,73
2,78	TERLAMBAT	2,94
3,09	TERLAMBAT	3,17
3,25	TERLAMBAT	3,28
3,3	TERLAMBAT	3,33
3,36	TEPAT WAKTU	3,36
3,36	TEPAT WAKTU	3,49
3,61	TEPAT WAKTU	3,68
3,75	TEPAT WAKTU	3,77
3,78	TEPAT WAKTU	

Tabel 21. Nilai *Threshold* Atribut IPS 5

IPS_5	SK	<i>Threshold</i>
0	DROP OUT	0,00
0	DROP OUT	0,00
0	DROP OUT	0,00
0	DROP OUT	0,00
0	DROP OUT	1,48
2,96	TERLAMBAT	3,14
3,32	TERLAMBAT	3,33
3,33	TERLAMBAT	3,36
3,38	TERLAMBAT	3,46
3,54	TEPAT WAKTU	3,54
3,54	TEPAT WAKTU	3,55
3,55	TERLAMBAT	3,63
3,71	TEPAT WAKTU	3,76
3,81	TEPAT WAKTU	3,82
3,82	TEPAT WAKTU	

Tabel 22. Nilai *Threshold* Atribut IPS 6

IPS_6	SK	<i>Threshold</i>
0	DROP OUT	0,00
0	DROP OUT	0,00
0	DROP OUT	0,00
0	DROP OUT	0,00
0	DROP OUT	1,38
2,75	TERLAMBAT	2,99
3,23	TERLAMBAT	3,31
3,39	TERLAMBAT	3,40
3,41	TEPAT WAKTU	3,43
3,44	TEPAT WAKTU	3,47
3,5	TEPAT WAKTU	3,57
3,64	TEPAT WAKTU	3,66
3,68	TEPAT WAKTU	3,68
3,68	TERLAMBAT	3,69
3,7	TERLAMBAT	

Tabel 23. Nilai *Threshold* Atribut Status Pernikahan

S_NIKAH	SK	<i>Threshold</i>
0	TEPAT WAKTU	0,0
0	TEPAT WAKTU	0,0
0	TERLAMBAT	0,0
0	TERLAMBAT	0,0
0	DROP OUT	0,0
0	TEPAT WAKTU	0,0
0	TEPAT WAKTU	0,0
0	DROP OUT	0,0
0	TERLAMBAT	0,0
0	TERLAMBAT	0,0
0	DROP OUT	0,0
0	TEPAT WAKTU	0,0
0	DROP OUT	0,5
1	DROP OUT	1,0
1	TERLAMBAT	

Tabel 24. Nilai *Threshold* Atribut Status Pekerjaan

S_KERJA	SK	<i>Threshold</i>
0	TEPAT WAKTU	0,0
0	TEPAT WAKTU	0,0
0	TERLAMBAT	0,0

S_KERJA	SK	Threshold
0	TERLAMBAT	0,0
0	DROP OUT	0,0
0	DROP OUT	0,0
0	TEPAT WAKTU	0,0
0	TEPAT WAKTU	0,0
0	DROP OUT	0,0
0	TERLAMBAT	0,0
0	TERLAMBAT	0,0
0	DROP OUT	0,0
0	TERLAMBAT	0,0
0	DROP OUT	0,5
1	TEPAT WAKTU	

Tabel 25. Nilai *Threshold* Atribut Jumlah SKS

J_SKS	SK	Threshold
19	DROP OUT	24,0
29	DROP OUT	33,5
38	DROP OUT	41,5
45	DROP OUT	62,5
80	DROP OUT	101,5
123	TERLAMBAT	124,5
126	TERLAMBAT	126,0
126	TERLAMBAT	126,5
127	TERLAMBAT	127,5
128	TEPAT WAKTU	128,5
129	TEPAT WAKTU	129,0
129	TEPAT WAKTU	129,5
130	TEPAT WAKTU	131,0
132	TERLAMBAT	132,5
133	TEPAT WAKTU	

Setelah mendapatkan nilai *threshold* masing-masing atribut, selanjutnya mencari nilai *Info* untuk masing-masing nilai *threshold* atribut menggunakan Persamaan 3:

Berikut nilai *Info* untuk nilai *threshold* atribut jenis kelamin <0

dan ≥ 0 :

$$\begin{aligned} Info(T) &= \left(-\left(\frac{0}{|0|}\right) \cdot \log_2 \left(\left(\frac{0}{|0|}\right) \right) + \left(-\left(\frac{0}{|0|}\right) \cdot \log_2 \left(\left(\frac{0}{|0|}\right) \right) + \left(-\left(\frac{0}{|0|}\right) \cdot \right. \right. \\ &\quad \left. \left. \log_2 \left(\left(\frac{0}{|0|}\right) \right) \right) \right) \\ &= 0 \end{aligned}$$

$$\begin{aligned} Info(T) &= \left(-\left(\frac{5}{|15|}\right) \cdot \log_2 \left(\left(\frac{5}{|15|}\right) \right) + \left(-\left(\frac{5}{|15|}\right) \cdot \log_2 \left(\left(\frac{5}{|15|}\right) \right) + \left(-\left(\frac{5}{|15|}\right) \cdot \right. \right. \\ &\quad \left. \left. \log_2 \left(\left(\frac{5}{|15|}\right) \right) \right) \right) \\ &= 1,584962501 \end{aligned}$$

Berikut nilai *Info* untuk nilai *threshold* atribut jenis kelamin

$<0,5$ dan $\geq 0,5$:

$$\begin{aligned} Info(T) &= \left(-\left(\frac{1}{|6|}\right) \cdot \log_2 \left(\left(\frac{1}{|6|}\right) \right) + \left(-\left(\frac{4}{|6|}\right) \cdot \log_2 \left(\left(\frac{4}{|6|}\right) \right) + \left(-\left(\frac{1}{|6|}\right) \cdot \right. \right. \\ &\quad \left. \left. \log_2 \left(\left(\frac{1}{|6|}\right) \right) \right) \right) \\ &= 1,251629 \end{aligned}$$

$$\begin{aligned} Info(T) &= \left(-\left(\frac{5}{|9|}\right) \cdot \log_2 \left(\left(\frac{5}{|9|}\right) \right) + \left(-\left(\frac{5}{|9|}\right) \cdot \log_2 \left(\left(\frac{5}{|9|}\right) \right) + \left(-\left(\frac{5}{|9|}\right) \cdot \right. \right. \\ &\quad \left. \left. \log_2 \left(\left(\frac{5}{|9|}\right) \right) \right) \right) \\ &= 1,392147 \end{aligned}$$

Berikut nilai *Info* untuk nilai *threshold* atribut jenis kelamin <1

dan ≥ 1 :

$$\begin{aligned} Info(T) &= \left(-\left(\frac{1}{|6|}\right) \cdot \log_2 \left(\left(\frac{1}{|6|}\right) \right) + \left(-\left(\frac{4}{|6|}\right) \cdot \log_2 \left(\left(\frac{4}{|6|}\right) \right) + \left(-\left(\frac{1}{|6|}\right) \cdot \right. \right. \\ &\quad \left. \left. \log_2 \left(\left(\frac{1}{|6|}\right) \right) \right) \right) \\ &= 1,251629 \end{aligned}$$

$$\begin{aligned} Info(T) &= \left(-\left(\frac{5}{|9|}\right) \cdot \log_2 \left(\left(\frac{5}{|9|}\right) \right) + \left(-\left(\frac{5}{|9|}\right) \cdot \log_2 \left(\left(\frac{5}{|9|}\right) \right) + \left(-\left(\frac{5}{|9|}\right) \cdot \right. \right. \\ &\quad \left. \left. \log_2 \left(\left(\frac{5}{|9|}\right) \right) \right) \right) \\ &= 1,392147 \end{aligned}$$

Setelah mendapatkan nilai *Info* dari masing-masing nilai *threshold* atribut, selanjutnya mencari nilai *Info* (*x*) menggunakan Persamaan:

$$Info_x(T) = \sum_{i=1}^n \left(\left(\frac{|Ti|}{|T|} \right) \cdot Info(Ti) \right) \quad (4)$$

Berikut nilai *Info*(*x*) untuk nilai *threshold* atribut jenis kelamin <0 dan >=0 :

$$\begin{aligned} Info(x) T &= \left(\left(\frac{|0|}{|15|} \right) \cdot 0 \right) + \left(\left(\frac{|15|}{|15|} \right) \cdot 1,584962501 \right) \\ &= 1,584962501 \end{aligned}$$

Berikut nilai *Info*(*x*) untuk nilai *threshold* atribut jenis kelamin <0,5 dan >=0,5 :

$$\begin{aligned} Info(x) T &= \left(\left(\frac{|6|}{|15|} \right) \cdot 1,251629 \right) + \left(\left(\frac{|9|}{|15|} \right) \cdot 1,392147 \right) \\ &= 1,33594 \end{aligned}$$

Berikut nilai *Info*(*x*) untuk nilai *threshold* atribut jenis kelamin <1 dan >=1 :

$$\begin{aligned} Info(x) T &= \left(\left(\frac{|6|}{|15|} \right) \cdot 1,251629 \right) + \left(\left(\frac{|9|}{|15|} \right) \cdot 1,392147 \right) \\ &= 1,33594 \end{aligned}$$

Setelah mendapatkan nilai *Info*(*x*) dari masing-masing nilai *threshold* atribut, selanjutnya mencari nilai *SplitInfo* menggunakan Persamaan :

$$SplitInfo(x) = - \sum_{i=1}^n \left(\left(\frac{|Ti|}{|T|} \right) \right) \cdot \log_2 \left(\left(\frac{|Ti|}{|T|} \right) \right) \quad (5)$$

Berikut nilai *SplitInfo* untuk nilai *threshold* atribut jenis kelamin <0 dan >=0 :

$$\begin{aligned} SplitInfo(x) &= \left(-\left(\frac{|0|}{|15|}\right)\right) \cdot \log_2 \left(\left(\frac{0}{|15|}\right)\right) + \left(-\left(\frac{15}{15}\right)\right) \cdot \log_2 \left(\left(\frac{15}{|15|}\right)\right) \\ &= 0 \end{aligned}$$

Berikut nilai *SplitInfo* untuk nilai *threshold* atribut jenis kelamin <0,5 dan >=0,5 :

$$\begin{aligned} SplitInfo(x) &= \left(-\left(\frac{|6|}{|15|}\right)\right) \cdot \log_2 \left(\left(\frac{6}{|15|}\right)\right) + \left(-\left(\frac{9}{15}\right)\right) \cdot \log_2 \left(\left(\frac{9}{|15|}\right)\right) \\ &= 0,970951 \end{aligned}$$

Berikut nilai *SplitInfo* untuk nilai *threshold* atribut jenis kelamin <1 dan >=1 :

$$\begin{aligned} SplitInfo(x) &= \left(-\left(\frac{|6|}{|15|}\right)\right) \cdot \log_2 \left(\left(\frac{6}{|15|}\right)\right) + \left(-\left(\frac{9}{15}\right)\right) \cdot \log_2 \left(\left(\frac{9}{|15|}\right)\right) \\ &= 0,970951 \end{aligned}$$

Setelah mendapatkan nilai *SplitInfo* dari masing-masing nilai *threshold* atribut, selanjutnya mencari nilai *Gain* menggunakan persamaan :

$$Gain(x) = Info T - Info x T \quad (6)$$

Berikut nilai *Gain* untuk nilai *threshold* atribut jenis kelamin <0 dan >=0 :

$$Gain(x) = 1,584962501 - 1,584962501 = 0$$

Berikut nilai *Gain* untuk nilai *threshold* atribut jenis kelamin <0,5 dan >=0,5 :

$$Gain(x) = 1,584962501 - 1,33594 = 0,249022$$

Berikut nilai *Gain* untuk nilai *threshold* atribut jenis kelamin <1 dan >=1 :

$$Gain(x) = 1,584962501 - 1,33594 = 0,249022$$

Setelah mendapatkan nilai *Gain* dari masing-masing nilai *threshold* atribut, selanjutnya mencari nilai *GainRatio* untuk menentukan nilai *threshold* yang akan digunakan sebagai nilai batas atribut, menggunakan Persamaan:

$$GainRatio = \frac{Gain(x)}{SplitInfo(x)} \quad (7)$$

Berikut nilai *GainRatio* untuk nilai *threshold* atribut jenis kelamin <0 dan >=0 :

$$GainRatio = \frac{0}{0} = 0$$

Berikut nilai *GainRatio* untuk nilai *threshold* atribut jenis kelamin <0,5 dan >=0,5 :

$$GainRatio = \frac{0,249022}{0,970951} = 0,256473$$

Berikut nilai *GainRatio* untuk nilai *threshold* atribut jenis kelamin <1 dan >=1 :

$$GainRatio = \frac{0,249022}{0,970951} = 0,256473$$

Berdasarkan nilai *GainRatio* terbesar maka nilai 0,5 atau 1 dapat dipilih sebagai *threshold* yang akan digunakan sebagai nilai batas atribut jenis kelamin.

Berikut nilai *Info* untuk nilai *threshold* atribut status pernikahan <0 dan >=0 :

$$\begin{aligned} Info(T) &= \left(-\left(\frac{0}{|0|}\right) \right) \cdot \log_2 \left(\left(\frac{0}{|0|}\right) \right) + \left(-\left(\frac{0}{|0|}\right) \right) \cdot \log_2 \left(\left(\frac{0}{|0|}\right) \right) + \\ &\quad \left(-\left(\frac{0}{|0|}\right) \right) \cdot \log_2 \left(\left(\frac{0}{|0|}\right) \right) \\ &= 0 \end{aligned}$$

$$\begin{aligned} Info(T) &= \left(-\left(\frac{5}{|15|}\right) \right) \cdot \log_2 \left(\left(\frac{5}{|15|}\right) \right) + \left(-\left(\frac{5}{|15|}\right) \right) \\ &\quad \log_2 \left(\left(\frac{5}{|15|}\right) \right) + \left(-\left(\frac{5}{|15|}\right) \right) \cdot \log_2 \left(\left(\frac{5}{|15|}\right) \right) \\ &= 1,584962501 \end{aligned}$$

Berikut nilai *Info* untuk nilai *threshold* atribut status pernikahan <0,5 dan >=0,5 :

$$\begin{aligned} Info(T) &= \left(-\left(\frac{5}{|13|}\right) \right) \cdot \log_2 \left(\left(\frac{5}{|13|}\right) \right) + \left(-\left(\frac{4}{|13|}\right) \right) \\ &\quad \cdot \log_2 \left(\left(\frac{4}{|13|}\right) \right) + \left(-\left(\frac{4}{|13|}\right) \right) \cdot \log_2 \left(\left(\frac{4}{|13|}\right) \right) \\ &= 1,576621 \end{aligned}$$

$$\begin{aligned} Info(T) &= \left(-\left(\frac{0}{|2|}\right) \right) \cdot \log_2 \left(\left(\frac{0}{|2|}\right) \right) + \left(-\left(\frac{1}{|2|}\right) \right) \\ &\quad \cdot \log_2 \left(\left(\frac{1}{|2|}\right) \right) + \left(-\left(\frac{1}{|2|}\right) \right) \cdot \log_2 \left(\left(\frac{1}{|2|}\right) \right) \\ &= 0 \end{aligned}$$

Berikut nilai *Info* untuk nilai *threshold* atribut status pernikahan <1 dan >=1 :

$$\begin{aligned} Info(T) &= \left(-\left(\frac{5}{|13|}\right) \right) \cdot \log_2 \left(\left(\frac{5}{|13|}\right) \right) + \left(-\left(\frac{4}{|13|}\right) \right) \cdot \log_2 \left(\left(\frac{4}{|13|}\right) \right) + \\ &\quad \left(-\left(\frac{4}{|13|}\right) \right) \cdot \log_2 \left(\left(\frac{4}{|13|}\right) \right) \end{aligned}$$

$$= 1,576621$$

$$\begin{aligned} Info(T) &= \left(-\left(\frac{0}{|2|}\right) \right) \cdot \log_2 \left(\left(\frac{0}{|2|}\right) \right) + \left(-\left(\frac{1}{|2|}\right) \right) \cdot \log_2 \left(\left(\frac{1}{|2|}\right) \right) + \\ &\quad \left(-\left(\frac{1}{|2|}\right) \right) \cdot \log_2 \left(\left(\frac{1}{|2|}\right) \right) \\ &= 0 \end{aligned}$$

Berikut nilai $Info(x)$ untuk nilai *threshold* atribut status pernikahan <0 dan ≥ 0 :

$$\begin{aligned} Info(x) T &= \left(\left(\frac{|0|}{|15|}\right) \cdot 0 \right) + \left(\left(\frac{|15|}{|15|}\right) \cdot 1,584962501 \right) \\ &= 1,584962501 \end{aligned}$$

Berikut nilai $Info(x)$ untuk nilai *threshold* atribut status pernikahan $<0,5$ dan $\geq 0,5$:

$$\begin{aligned} Info(x) T &= \left(\left(\frac{|13|}{|15|}\right) \cdot 1,576621 \right) + \left(\left(\frac{|2|}{|15|}\right) \cdot 0 \right) \\ &= 1,366405 \end{aligned}$$

Berikut nilai $Info(x)$ untuk nilai *threshold* atribut status pernikahan <1 dan ≥ 1 :

$$\begin{aligned} Info(x) T &= \left(\left(\frac{|13|}{|15|}\right) \cdot 1,576621 \right) + \left(\left(\frac{|2|}{|15|}\right) \cdot 0 \right) \\ &= 1,366405 \end{aligned}$$

Berikut nilai $SplitInfo$ untuk nilai *threshold* atribut status pernikahan <0 dan ≥ 0 :

$$\begin{aligned}
SplitInfo(x) &= \left(-\left(\frac{|0|}{|15|}\right) \right) \cdot \log_2 \left(\left(\frac{0}{|15|}\right) \right) + \left(-\left(\frac{15}{15}\right) \right) \cdot \\
&\quad \log_2 \left(\left(\frac{15}{|15|}\right) \right) \\
&= 0
\end{aligned}$$

Berikut nilai *SplitInfo* untuk nilai *threshold* atribut status pernikahan <0,5 dan >=0,5 :

$$\begin{aligned}
SplitInfo(x) &= \left(-\left(\frac{|13|}{|15|}\right) \right) \cdot \log_2 \left(\left(\frac{13}{|15|}\right) \right) + \left(-\left(\frac{2}{15}\right) \right) \cdot \\
&\quad \log_2 \left(\left(\frac{2}{|15|}\right) \right) \\
&= 0,56651
\end{aligned}$$

Berikut nilai *SplitInfo* untuk nilai *threshold* atribut status pernikahan <1 dan >=1 :

$$\begin{aligned}
SplitInfo(x) &= \left(-\left(\frac{|13|}{|15|}\right) \right) \cdot \log_2 \left(\left(\frac{13}{|15|}\right) \right) + \left(-\left(\frac{2}{15}\right) \right) \cdot \\
&\quad \log_2 \left(\left(\frac{2}{|15|}\right) \right) \\
&= 0,56651
\end{aligned}$$

Berikut nilai *Gain* untuk nilai *threshold* atribut status pernikahan <0 dan >=0 :

$$Gain(x) = 1,584962501 - 1,584962501 = 0$$

Berikut nilai *Gain* untuk nilai *threshold* atribut status pernikahan <0,5 dan >=0,5 :

$$Gain(x) = 1,584962501 - 1,366405 = 0,218557$$

Berikut nilai *Gain* untuk nilai *threshold* atribut status pernikahan <1 dan >=1 :

$$Gain(x) = 1,584962501 - 1,33594 = 0,249022$$

Berikut nilai *GainRatio* untuk nilai *threshold* atribut status pernikahan <0 dan >=0 :

$$GainRatio = \frac{0}{0} = 0$$

Berikut nilai *GainRatio* untuk nilai *threshold* atribut status pernikahan <0,5 dan >=0,5 :

$$GainRatio = \frac{0,218557}{0,56651} = 0,385797$$

Berikut nilai *GainRatio* untuk nilai *threshold* atribut status pernikahan <1 dan >=1 :

$$GainRatio = \frac{0,249022}{0,970951} = 0,256473$$

Berdasarkan nilai *GainRatio* terbesar maka nilai 0,5 dipilih sebagai *threshold* yang akan digunakan sebagai nilai batas atribut status pernikahan.

Berikut nilai *Info* untuk nilai *threshold* atribut status pekerjaan <0 dan >=0 :

$$\begin{aligned} Info(T) &= \left(-\left(\frac{0}{|0|} \right) \right) \cdot \log_2 \left(\left(\frac{0}{0} \right) \right) + \left(-\left(\frac{0}{|0|} \right) \right) \\ &\quad \cdot \log_2 \left(\left(\frac{0}{|0|} \right) \right) + \left(-\left(\frac{0}{|0|} \right) \right) \cdot \log_2 \left(\left(\frac{0}{|0|} \right) \right) \\ &= 0 \end{aligned}$$

$$\begin{aligned}
Info(T) &= \left(-\left(\frac{5}{|15|}\right) \right) \cdot \log_2 \left(\left(\frac{5}{|15|}\right) \right) + \left(-\left(\frac{5}{|15|}\right) \right) \\
&\quad \log_2 \left(\left(\frac{5}{|5|}\right) \right) + \left(-\left(\frac{5}{|5|}\right) \right) \cdot \log_2 \left(\left(\frac{5}{|5|}\right) \right) \\
&= 1,584963
\end{aligned}$$

Berikut nilai *Info* untuk nilai *threshold* atribut status pekerjaan

<0,5 dan >=0,5 :

$$\begin{aligned}
Info(T) &= \left(-\left(\frac{5}{|13|}\right) \right) \cdot \log_2 \left(\left(\frac{5}{|13|}\right) \right) + \left(-\left(\frac{4}{|13|}\right) \right) \\
&\quad \cdot \log_2 \left(\left(\frac{4}{|13|}\right) \right) + \left(-\left(\frac{4}{|13|}\right) \right) \cdot \log_2 \left(\left(\frac{4}{|13|}\right) \right) \\
&= 1,576621
\end{aligned}$$

$$\begin{aligned}
Info(T) &= \left(-\left(\frac{0}{|2|}\right) \right) \cdot \log_2 \left(\left(\frac{0}{|2|}\right) \right) + \left(-\left(\frac{1}{|2|}\right) \right) \\
&\quad \cdot \log_2 \left(\left(\frac{1}{|2|}\right) \right) + \left(-\left(\frac{1}{|2|}\right) \right) \cdot \log_2 \left(\left(\frac{1}{|2|}\right) \right) \\
&= 0
\end{aligned}$$

Berikut nilai *Info(x)* untuk nilai *threshold* atribut status pekerjaan <0 dan >=0 :

$$\begin{aligned}
Info(x) T &= \left(\left(\frac{|0|}{|15|}\right) \cdot 0 \right) + \left(\left(\frac{|15|}{|15|}\right) \cdot 1,584963 \right) \\
&= 0
\end{aligned}$$

Berikut nilai *Info(x)* untuk nilai *threshold* atribut status pekerjaan <0,5 dan >=0,5 :

$$\begin{aligned}
Info(x) T &= \left(\left(\frac{|13|}{|15|}\right) \cdot 1,576621 \right) + \left(\left(\frac{|2|}{|15|}\right) \cdot 0 \right) \\
&= 1,366405
\end{aligned}$$

Berikut nilai *SplitInfo* untuk nilai *threshold* atribut status pekerjaan <0 dan >=0 :

$$\begin{aligned}
 SplitInfo(x) &= \left(-\left(\frac{|0|}{|15|} \right) \right) \cdot \log_2 \left(\left(\frac{0}{|15|} \right) \right) + \left(-\left(\frac{15}{15} \right) \right) \cdot \\
 &\quad \log_2 \left(\left(\frac{15}{|15|} \right) \right) \\
 &= 0
 \end{aligned}$$

Berikut nilai *SplitInfo* untuk nilai *threshold* atribut status pekerjaan <0,5 dan >=0,5 :

$$\begin{aligned}
 SplitInfo(x) &= \left(-\left(\frac{|14|}{|15|} \right) \right) \cdot \log_2 \left(\left(\frac{14}{|15|} \right) \right) + \left(-\left(\frac{1}{15} \right) \right) \cdot \\
 &\quad \log_2 \left(\left(\frac{1}{|15|} \right) \right) \\
 &= 0,353359
 \end{aligned}$$

Berikut nilai *Gain* untuk nilai *threshold* atribut status pekerjaan <0 dan >=0 :

$$Gain(x) = 1,584962501 - 0 = 0$$

Berikut nilai *Gain* untuk nilai *threshold* atribut status pekerjaan <0,5 dan >=0,5 :

$$Gain(x) = 1,584962501 - 1,472246 = 0,112717$$

Berikut nilai *GainRatio* untuk nilai *threshold* atribut status pekerjaan <0 dan >=0 :

$$GainRatio = \frac{0}{0} = 0$$

Berikut nilai *GainRatio* untuk nilai *threshold* atribut status pekerjaan <0,5 dan >=0,5 :

$$GainRatio = \frac{0,112717}{0,353359} = 0,318986$$

Berdasarkan nilai *GainRatio* terbesar maka nilai 0,5 dipilih sebagai *threshold* yang akan digunakan sebagai nilai batas atribut status pekerjaan.

Tabel 26. Nilai *Threshold* Berdasarkan *GainRatio* Terbesar

Atribut	GainRatio
Jenis Kelamin	0,5
IPS 1	2,62
IPS 2	2,48
IPS 3	2,62
IPS 4	3,33
IPS 5	1,48
IPS 6	1,38
Status Pernikahan	0,5
Status Pekerjaan	0,5
Jumlah SKS	101,5

Tabel 26 merupakan nilai *threshold* berdasarkan nilai *GainRatio* terbesar. Nilai-nilai tersebut digunakan sebagai nilai batas atribut untuk pelatihan seperti pada Tabel 27.

Tabel 27. Tabel Perhitungan Algoritma C4.5

NODE	ATRIBUT	NILAI	TOTAL	TEPAT WAKTU	TERLAMBAT	DROP OUT	ENTROPY	GAIN
1	Jenis Kelamin	TOTAL						
		<0,5						
		>=0,5						
	IPS 1							
		<2,62						
		>=2,62						

NODE	ATRIBUT	NILAI	TOTAL	TEPAT WAKTU	TERLAMBAT	DROP OUT	ENTROPY	GAIN
1	IPS 2	<2,48						
		>=2,48						
	IPS 3	<2,62						
		>=2,62						
	IPS 4	<3,33						
		>=3,33						
	IPS 5	<1,48						
		>=1,48						
	IPS 6	<1,38						
		>=1,38						
	Status Pernikahan	<0,5						
		>=0,5						
	Status Pekerjaan	<0,5						
		>=0,5						
	Jumlah SKS	<101,5						
		>=101,5						

Langkah pertama perhitungan dengan algoritma C4.5 yaitu menghitung *entropy* keseluruhan kasus menggunakan Persamaan 2.

$$\begin{aligned}
 Entropy &= \left(-\left(\frac{5}{15}\right) \right) \cdot \log_2 \left(\left(\frac{5}{15}\right) \right) + \left(-\left(\frac{5}{15}\right) \right) \cdot \log_2 \left(\left(\frac{5}{15}\right) \right) + \\
 &\left(-\left(\frac{5}{15}\right) \right) \cdot \log_2 \left(\left(\frac{5}{15}\right) \right) \\
 &= 1,584962501
 \end{aligned}$$

Setelah mendapatkan *entropy* keseluruhan data, selanjutnya menghitung *entropy* dari masing-masing atribut menggunakan Persamaan 2.

Tabel 28. Hasil Perhitungan *Entropy*

NODE	ATRIBUT	NILAI	TOTAL	TEPAT WAKTU	TERLAMBAT	DROP OUT	ENTROPY	GAIN
1		TOTAL	15	5	5	5		
							1,584963	
	Jenis Kelamin	<0,5	6	1	4	1	1,251629	
		>=0,5	9	4	1	4	1,392147	
	IPS 1	<2,62	5	0	0	5	0	
		>=2,62	10	5	5	0	0	
	IPS 2	<2,48	5	0	0	5	0	
		>=2,48	10	5	5	0	0	
	IPS 3	<2,62	5	0	0	5	0	
		>=2,62	10	5	5	0	0	
	IPS 4	<3,33	10	0	5	5	0	
		>=3,33	5	5	0	0	0	
	IPS 5	<1,48	5	0	0	5	0	
		>=1,48	10	5	5	0	0	
	IPS 6	<1,38	5	0	0	5	0	
		>=1,38	10	5	5	0	0	
	Status Pernikahan	<0,5	13	5	4	4	1,576621	
		>=0,5	2	0	1	1	0	
	Status Pekerjaan	<0,5	14	4	5	5	1,577406	
		>=0,5	1	1	0	0	0	
	Jumlah SKS	<101,5	5	0	0	5	0	
		>=101,5	10	5	5	0	0	

Setelah mendapatkan *entropy* masing-masing atribut, selanjutnya menghitung nilai *gain* menggunakan Persamaan 1.

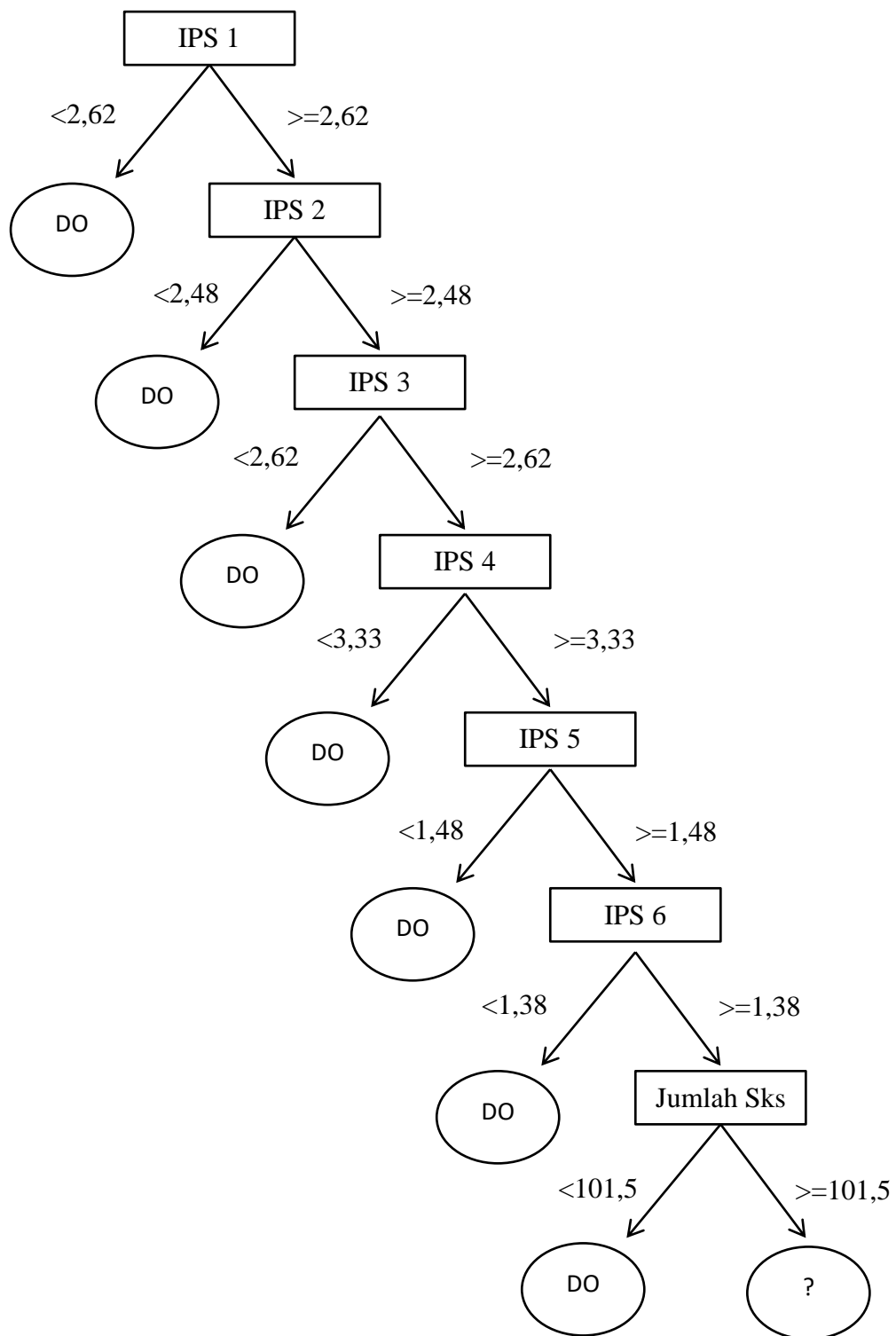
$$Gain = 1,584963 - \left(\left(\frac{|6|}{|15|} \cdot 1,251629 \right) + \left(\frac{|9|}{|15|} \right) \cdot 1,392147 \right) \\ = 0,249022$$

Tabel 29. Hasil Perhitungan *Gain*

NODE	ATRIBUT	NILAI	TOTAL	TEPAT WAKTU	TERLAMBAT	DROP OUT	ENTROPY	GAIN
1		TOTAL	15	5	5	5		
							1,584963	
	Jenis Kelamin	<0,5	6	1	4	1	1,251629	
		>=0,5	9	4	1	4	1,392147	
								0,249022
	IPS 1	<2,62	5	0	0	5	0	
		>=2,62	10	5	5	0	0	
								1,584963
	IPS 2	<2,48	5	0	0	5	0	
		>=2,48	10	5	5	0	0	
								1,584963
	IPS3	<2,62	5	0	0	5	0	
		>=2,62	10	5	5	0	0	
								1,584963
	IPS 4	<3,33	10	0	5	5	0	
		>=3,33	5	5	0	0	0	
								1,584963
	IPS 5	<1,48	5	0	0	5	0	
		>=1,48	10	5	5	0	0	
								1,584963
	IPS 6	<1,38	5	0	0	5	0	
		>=1,38	10	5	5	0	0	
								1,584963
	Status Pernikahan	<0,5	13	5	4	4	1,576621	
		>=0,5	2	0	1	1	0	
								0,218557
	Status Pekerjaan	<0,5	14	4	5	5	1,577406	
		>=0,5	1	1	0	0	0	
								0,112717

NOD E	ATRIBUT	NILAI	TOTAL	TEPAT WAKTU	TERLAMBAT	DROP OUT	<i>ENTROPY</i>	<i>GAIN</i>
	Jumlah SKS	<101,5	5	0	0	5	0	
		>=101,5	10	5	5	0	0	
								1,584963

Berdasarkan Tabel 29 *Gain* terbesar ada pada atribut IPS 1, IPS 2, IPS 3, IPS 4, IPS 5, IPS 6 dan jumlah SKS.



Gambar 27. Pohon Keputusan Node 1

Gambar 27 merupakan pohon keputusan node 1. Pada pohon keputusan IPS 1 <2,62 memiliki 5 data dan semua termasuk dalam status kelulusan *Drop Out*. Maka IPS 1 <2,62 menjadi daun atau *leaf*. Demikian dengan IPS 2, IPS 3, IPS 4, IPS 5, IPS 6 dan jumlah SKS. Pohon keputusan node 1 memiliki banyak cabang yang mencerminkan adanya *noise* pada pelatihan data. Oleh karena itu pohon node 1 dipangkas menggunakan Persamaan :

$$e = \frac{\left(f + \frac{z^2}{2N} + z \sqrt{\frac{f}{N} - \frac{f^2}{N} + \frac{z^2}{4N^2}}\right)}{\left(1 + \frac{z^2}{N}\right)} \quad (8)$$

Keterangan :

f : nilai perbandingan *error rate*

N : total data

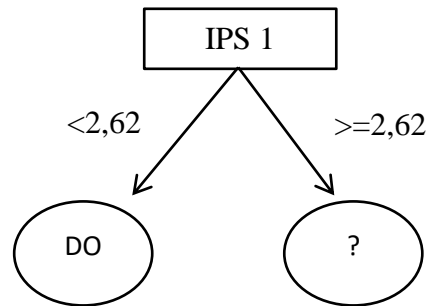
Z : *confidence level* algoritma C4.5 = 25% = 0,69 berdasarkan tabel distribusi normal

$$f = \frac{0}{15} = 0$$

$$N = 15$$

$$\begin{aligned} e &= \frac{\left(0 + \frac{0,69^2}{2 \cdot 15} + 0,69 \sqrt{\frac{0}{15} - \frac{0^2}{15} + \frac{0,69^2}{4 \cdot 15^2}}\right)}{\left(1 + \frac{0,69^2}{15}\right)} \\ &= \frac{0,01587 + 0,0614645}{0,03174} \\ &= 2,43649968 \\ &= 0,024365\% \end{aligned}$$

Dengan demikian tingkat *error* pada pohon keputusan setelah dipangkas adalah sebesar 0,024365%.



Gambar 28. Pohon Keputusan Node 1 Setelah Dipangkas

Berdasarkan Gambar 28 selanjutnya akan dianalisis node 1.1 dengan data IPS 1 $\geq 2,62$.

Tabel 30. Data yang Memiliki IPS 1 $\geq 2,62$

JK	IPS_1	IPS_2	IPS_3	IPS_4	IPS_5	IPS_6	S_NIKAH	S_KERJA	J_SKS	SK
1	3,32	3,45	3,55	3,78	3,71	3,41	0	0	129	TEPAT WAKTU
1	3,26	3,18	3,55	3,61	3,81	3,68	0	0	129	TEPAT WAKTU
1	2,74	2,60	2,80	2,67	3,33	3,23	0	0	123	TERLAMBAT
0	2,84	2,85	3	2,78	3,38	3,68	0	0	127	TERLAMBAT
1	2,76	2,71	2,91	3,36	3,54	3,44	0	0	128	TEPAT WAKTU
1	2,95	3,05	3,08	3,36	3,82	3,5	0	0	130	TEPAT WAKTU
0	2,67	2,57	2,83	3,09	2,96	2,75	0	0	132	TERLAMBAT
0	3,05	2,74	2,82	3,25	3,55	3,39	0	0	126	TERLAMBAT
0	2,86	2,86	3,39	3,75	3,54	3,64	0	1	133	TEPAT WAKTU
0	3,38	3,13	3	3,3	3,32	3,7	1	0	126	TERLAMBAT

Tabel 30 merupakan data yang memiliki IPS 1 $\geq 2,62$. Data tersebut dianalisis menggunakan Persamaan 1 dan Persamaan 2.

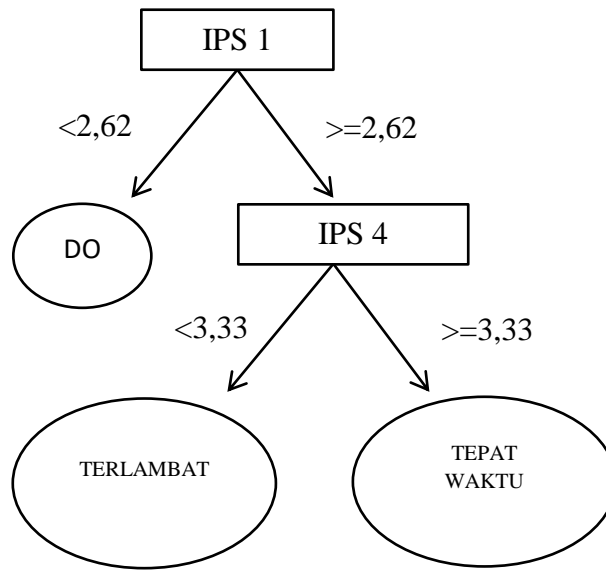
Hitung *entropy* dari atribut IPS 1 $\geq 2,62$:

$$\begin{aligned}
 Entropy &= \left(-\left(\frac{5}{|S|} \right) \right) \cdot \log_2 \left(\left(\frac{5}{|S|} \right) \right) + \left(-\left(\frac{5}{|S|} \right) \right) \cdot \log_2 \left(\left(\frac{5}{|S|} \right) \right) + \\
 &\quad \left(-\left(\frac{0}{|S|} \right) \right) \cdot \log_2 \left(\left(\frac{0}{|S|} \right) \right) = 0
 \end{aligned}$$

Setelah mendapatkan *entropy* dari atribut IPS $\geq 2,62$,
selanjutnya hitung *entropy* masing-masing atribut menggunakan
Persamaan 2 dan *gain*nya menggunakan Persamaan 1.

Tabel 31. Hasil Perhitungan Node 1.1

NODE	ATRIBUT	NILAI	TOTAL	TEPAT WAKTU	TERLAMBAT	DROP OUT	ENTROPY	GAIN
1,1		TOTAL	10	5	5	0		
							0	
	Jenis Kelamin	<0,5	5	1	4	0		
		$\geq 0,5$	5	4	1	0		
							0	0
	IPS2	<2,48	0	0	0	0		
		$\geq 2,48$	10	5	5	0		
							0	0
	IPS 3	<2,62	0	0	0	0		
		$\geq 2,62$	10	5	5	0		
							0	0
	IPS 4	<3,33	5	0	5	0		
		$\geq 3,33$	5	5	0	0		
							0	0
	IPS 5	<1,48	0	0	0	0		
		$\geq 1,48$	10	5	5	0		
							0	0
	IPS 6	<1,38	0	0	0	0		
		$\geq 1,38$	10	5	5	0		
							0	0
	Status Pernikahan	<0,5	9	5	4	0		
		$\geq 0,5$	1	0	1	0		
							0	0
	Status Pekerjaan	<0,5	9	4	5	0		
		$\geq 0,5$	1	1	0	0		
							0	0
	Jumlah SKS	<101,5	0	0	0	0		
		$\geq 101,5$	10	5	5	0		
							0	0



Gambar 29. Pohon Keputusan Akhir

Gambar 29 merupakan pohon keputusan akhir karena semua data sudah mempunyai kelas. Atribut IPS 4 diambil karena IPS 4 $<3,33$ memiliki 5 data dengan keluaran terlambat semua dan $\geq 3,33$ memiliki 5 data dengan keluaran tepat waktu semua.

Berdasarkan pohon keputusan akhir akan dilakukan pengujian data berikut :

NIM	: C1455201021
Nama	: Pebriansyah
Jenis Kelamin	: Laki-Laki
IPS 1	: 2,95
IPS 2	: 2,86
IPS 3	: 2,70
IPS 4	: 3,27
IPS 5	: 3,09
IPS 6	: 3,57
Status Pernikahan	: Belum
Status Pekerjaan	: Belum
Jumlah SKS	: 127

Berdasarkan pohon keputusan pada Gambar 28 data tersebut termasuk dalam *output* tepat waktu.

Gambar 30. Hasil Prediksi Sistem

Gambar 30 menunjukkan bahwa hasil prediksi yang dilakukan oleh sistem dan hasil prediksi berdasarkan pohon keputusan Gambar 29 memiliki *output* yang sama.

Tabel 32. Perbandingan Pengujian Data

Data Latih	89	89	60	60
Data Uji	31	20	31	20
Sesuai	23	14	12	7
Tidak Sesuai	8	6	19	13

Berdasarkan Tabel 32 maka didapat persentase perbandingan yaitu seperti pada Tabel 33.

Tabel 33. Persentase Perbandingan Data

Data Latih	89	89	60	60
Data Uji	31	20	31	20
Sesuai	74,19%	70%	38,71%	35%
Tidak Sesuai	25,81%	30%	61,29%	65%

Berdasarkan Tabel 33 diketahui bahwa pengujian dengan data latih sebanyak 89 data menghasilkan akurasi yang lebih besar dibandingkan dengan pengujian data latih sebanyak 60 data. Pengujian dengan data uji sebanyak 31 data memiliki akurasi tidak sesuai lebih kecil dibandingkan dengan pengujian data uji sebanyak 20 data.

BAB V

KESIMPULAN DAN SARAN

A. Kesimpulan

Berdasarkan hasil penelitian maka dapat diambil kesimpulan sebagai berikut :

1. Implementasi algoritma C4.5 dalam prediksi kelulusan mahasiswa STMIK Palangkaraya menggunakan *software* Matlab.
2. Metode *data mining* dengan algoritma C4.5 dapat diimplementasikan dalam prediksi kelulusan mahasiswa STMIK Palangkaraya.
3. Berdasarkan hasil uji coba, semakin banyak data yang digunakan maka semakin besar tingkat akurasi algoritma.

B. Saran

Beberapa saran yang dapat dijadikan sebagai pertimbangan untuk penelitian ke depan :

1. Menambahkan beberapa atribut dalam melakukan prediksi untuk mendapatkan hasil prediksi yang lebih baik.
2. Melakukan prediksi kelulusan mahasiswa dengan menggabungkan algoritma C4.5 dengan metode yang lain untuk mendapatkan nilai akurasi yang lebih baik.
3. Menambahkan *output* sistem berupa peluang kelulusan dengan nilai IPK sesuai dengan nilai standar kelulusan.

DAFTAR PUSTAKA

- Adelia & Jimmy Setiawan, 2011, *Implementasi Customer Relationship Management (CRM) pada Sistem Reservasi Hotel Berbasis Website dan Dekstop*, Sistem Informasi, Vol. VI, Nomor 02
- Azwanti, Nurul, 2018, *Analisis Algoritma C4.5 untuk Memprediksi Penjualan Motor pada PT. Capella Dinamik Nusantara Cabang Muka Kuning*, Ilmu Komputer, Vol. XIII, Nomor 01
- Harahap, Fitriana, 2015, *Penerapan Data Mining dalam Prediksi Pembelian Cat*, Konferensi Nasional Sistem & Informasi
- Hermanto, Bambang, Azhari SN & Fajri Profesio Putra, 2017, *Analisis Kinerja Decision Tree dalam Prediksi Potensi Pelunasan Kredit Calon Debitur*, Inovtek Polbeng, Vol. II, Nomor 02
- Maharani, Mega & Nita Merlina, 2014, *Penerapan Metode Straight Selection pada Sistem Parkir Universitas Bina Nusantara*, Pilar Nusa Mandiri, Vol. X, Nomor 01
- Oktavia, Nia & Hutrianto, 2016, *Extreme Programming sebagai Metode Pengembangan E-Keuangan pada Pondok Pesantren Qodratullah*, Ilmiah Matrix, Vol. XVIII, Nomor 02
- Prasetyo, Eko. 2014. *Data Mining : Mengolah Data Menjadi Informasi Menggunakan Matlab*, Andi, Yogyakarta
- Putri, Ratna Puspita Sari & Indra Waspada, 2018, *Penerapan Algoritma C4.5 pada Aplikasi Prediksi Kelulusan Mahasiswa Prodi Informatika*, Ilmu Komputer dan Informatika, Vol. IV, Nomor 01
- RM, Bagus Winahyu. 2016. *Belajar Mudah dan Cepat Microsoft Excel 2010*, Pustaka Baru Press, Yogyakarta.
- Romadhona, Agus, Suprapedi & H. Himawan, 2017, *Prediksi Kelulusan Mahasiswa Tepat Waktu Berdasarkan Usia, Jenis Kelamin dan Indeks Prestasi Menggunakan Algoritma Desicion Tree*, Teknologi Informasi, Vol. XIII, Nomor 01
- Rosandy, Triowali, 2016, *Perbandingan Metode Naive Bayes Classifier dengan Metode Desicion Tree (C4.5) untuk Menganalisis Kelancaran Pembiayaan (Studi Kasus: KSPPS/BMT AL-FADHILA)*, Tim Darmajaya, Vol. II, Nomor 01

Sidi, M. Mustaqbal. Roeri Fajri Firdaus & HendraRahmadi, *Pengujian Aplikasi Menggunakan Black Box Testing Boundary Value Analysis*, Vol. I, Nomor 03

Sitorus, L. 2015. *Algoritma dan Pemrograman*, Andi, Yogyakarta

STMIK Palangkaraya, 2018. *Pedoman Penulisan Proposal dan Tugas Akhir Teknik Informatika 2018*. Palangkaraya

Sudoryono. 2015. *Metodologi Riset di Bidang IT : Panduan Praktis, Teori dan Contoh Kasus*, ANDI, Yogyakarta

Supriyatna, Adi, 2013, *Metode Extreme Programming Pada Pembangunan WEB Aplikasi Seleksi Peserta Pelatihan Kerja*, Teknik Informatika, Vol. XI, Nomor 01

Tjolleng, Amir. 2017. *Pengantar Pemograman MATLAB*, PT. Elex Media Komputindo, Jakarta

Widodo, Herlawati, Rahmadya Trias Handayanto. 2013. *Penerapan Data Mining dengan Matlab*, Informatika, Makassar

LAMPIRAN

Kepada :

Yth. Ketua STMIK Palangkaraya

Jl. G.Obos N0. 114

di-

Palangkaraya

Dengan hormat,

Sehubungan dengan penyusunan skripsi mahasiswa sebagai persyaratan kelulusan Program Strata I (S-1) pada Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) Palangkaraya, maka bersama ini saya sampaikan permohonan ijin penelitian untuk :

- | | |
|--------------------------|---|
| 1. Nama Lengkap | : Teti Sumarni |
| 2. NIM | : C1555201023 |
| 3. Program/Jurusan | : Teknik Informatika |
| 4. Semester-Th. Akademik | : VIII-2018/2019 |
| 5. Alamat | : Tanjung Karitak |
| 6. Waktu/Lama Penelitian | : 16 Mei 2019- 16 Juni 2019 |
| 7. Tempat Penelitian | : STMIK Palangkaraya/ Unit AKMA
(Akademik Kemahasiswaan) |

Dengan judul skripsi :

**“IMPLEMENTASI ALGORITMA C4.5 DALAM PREDIKSI KELULUSAN
MAHASISWA STMIK PALANGKARAYA”**

Adapun ketentuan dan tata tertib pemberian informasi dan data yang diperlukan dalam penelitian tersebut menyesuaikan dengan ketentuan/peraturan yang berlaku.

Demikian permohonan ini saya sampaikan, atas bantuan dan kerjasamanya diucapkan terimakasih.

Palangkaraya, 16 Mei 2019

Ybs



Teti Sumarni



**SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
(STMIK) PALANGKARAYA**

Jl. G. Obos No.114 Telp. 0536-3224593, 3225515 Fax. 0536-3225515 Palangkaraya
surel (email) : humas@stmikplk.ac.id – laman (website) : www.stmikplk.ac.id

SURAT TUGAS PENGUJI TUGAS AKHIR

No. 351 /STMIK-3.C.2/AK/VII/2019

Ketua Program Studi Teknik Informatika Sekolah Tinggi Manajemen Informatika dan Komputer
(STMIK) Palangkaraya menugaskan kepada nama- nama berikut :

1. Nama : **Frengklin Matatula,S.Kom., M.Msi**
NIK : 198611012015102
Sebagai Ketua
2. Nama : **Arliyana, M.Kom**
NIK : 198510082010002
Sebagai Sekretaris
3. Nama : **Sam'ani, ST., M.Kom**
NIK : 197703252005105
Sebagai Anggota
4. Nama : **Lili Rusdiana, M.Kom**
NIK : 198707282011007
Sebagai Anggota
5. Nama : **Elok Faiqotul Himmah, S.Si., M.Sc**
NIK : 198503092009003
Sebagai Anggota

Tim Penguji Tugas Akhir Mahasiswa :

- Nama : **Teti Sumarni**
NIM : C1555201023
Hari/ Tanggal Ujian : **Kamis, 25 Juli 2019**
Waktu : **08.30 WIB - Selesai**
Judul Tugas Akhir : Implementasi Algoritma C4.5 Dalam Prediksi Kelulusan
Mahasiswa STMIK Palangka Raya

Demikian surat ini dibuat agar dapat dipergunakan sebagaimana mestinya dan dilaksanakan
dengan penuh tanggung jawab.

Palangka Raya, 24 Juli 2019

Ketua Program Studi,

Hotmian Sitohang, M.Kom.
NIK. 198503282008002



SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
(STMIK) PALANGKARAYA

Jl. G. Obos No.114 Telp. 0536-3224593 Fax. 0536-3225515 Palangkaraya
surel (email) : humas@stmikplk.ac.id – laman (website) : www.stmikplk.ac.id

SURAT TUGAS

No. 179/STMIK-3.C.2/AU/II/2019

Ketua Program Studi Teknik Informatika Sekolah Tinggi Manajemen Informatika dan Komputer
(STMIK) Palangkaraya menugaskan nama-nama tersebut di bawah ini :

Nama : Lili Rusdiana, M.Kom

NIK : 198707282011007

Sebagai Pembimbing I dalam Pembuatan Program

Nama : Elok Faiqotul Himmah, S.Si, M.Sc

NIK : 198503092009003

Sebagai Pembimbing II dalam Penulisan Tugas Akhir

Untuk membimbing Tugas Akhir mahasiswa :

Nama : Teti Sumarni

NIM : C1555201023

Program Studi: TEKNIK INFORMATIKA (55201)

Tanggal Daftar: 16 Januari 2019

Judul Skripsi : Implementasi Algoritma C45 Dalam Prediksi Mahasiswa STMIK
PALANGKARAYA

Demikian surat ini dibuat agar dapat dipergunakan sebagaimana mestinya dan dilaksanakan
dengan penuh tanggung jawab.

Palangkaraya, 16 Januari 2019

Ketua Program Studi Teknik Informatika

Hotmian Sitohang, M.Kom
NIK. 198503282008002

Tembusan :

1. Kepala Unit Penjaminan Mutu Internal (UPMI) STMIK Palangkaraya



SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
(STMIK) PALANGKARAYA

Jl. G. Obos No.114 Telp.0536-3224593, 3225515 Fax.0536-3225515 Palangkaraya
email : humas@stmikplk.ac.id - website : www.stmikplk.ac.id

KARTU KEGIATAN KONSULTASI
TUGAS AKHIR

NIM - Nama Lengkap : C1555201023 - TETI SUMARNI
Nomor Telpn / HP : 082113858596
Dosen Pembimbing : 1. Lili Rusdiana, M.Kom
2. Elok Faiqotul Himmah, S.Si., M.Sc.,
Tanggal Surat Tugas : 16/01/2019
Judul Tugas Akhir : Implementasi Algoritma C4.5 Dalam Prediksi Kelulusan Mahasiswa
STMIK Palangkaraya

No.	Tanggal Konsultasi		Catatan Bimbingan	Tanda Tangan
	Terima	Kembali		
	28/02-2019	04/03-2019	Perbaiki yang ditandai sesuai catatan	Ajij
	06/02-2019	08/03-2019	Perbaiki yg ditandai	Ajij
	13/3-2019	13/3-2019	lengkapi penulisan dan pendukung metode pada penulisan	Ajij
	19/3-2019	19/3-2019	Papitan penulisan & Ace seminar	Ajij
	21/3-2019	22/3-2019	Cantumkan penulis referensi yg digunakan di Zeffar pustak ace seminar	Ajij
	9/5-2019	11/5-2019	Perbaiki analisis kelemahan sistem & penulisan yg ditandai	Ajij
	27/6-2019	27/6-2019	Penyesuaian Data	Ajij
	5/7-2019	5/7-2019	Sesuaikan pada batasan masalah	Ajij
	13/7-2019	13/7-2019	Ace program, lanjut uji coba & penulisan	Ajij
	16/7-2019	16/7-2019	Lanjutan ke bab berikutnya	Ajij
	22/7-2019		Perbaiki perhitungan yg msl salah	Ajij

Palangkaraya, 28/02/2019

Tanggal Daftar Ujian :

Mahasiswa Ybs,

Ajij

Lampiran 2. Data Latih

NIM	NAMA	JK	IPS_1	IPS_2	IPS_3	IPS_4	IPS_5	IPS_6	S_NIKAH	S_KERJA	J_SKS	SK
C1255201001	ROBY SETIAWAN	L	3,32	3,45	3,55	3,78	3,71	3,41	BELUM	BELUM	129	TEPAT WAKTU
C1255201002	GUSTRIYANA DEWI	P	3,61	2,86	3,00	3,35	3,57	3,68	BELUM	BELUM	129	TERLAMBAT
C1255201003	RADEN RIZKY PINASTI	L	3,26	3,18	3,55	3,61	3,81	3,68	BELUM	BELUM	129	TEPAT WAKTU
C1255201005	EDY KURNIAWAN	L	2,74	2,60	2,80	2,67	3,33	3,23	BELUM	BELUM	123	TERLAMBAT
C1255201006	YULIANA PERMATA SARI	P	2,84	2,85	3	2,78	3,38	3,68	BELUM	BELUM	127	TERLAMBAT
C1255201007	PEDTRIYANA SHISILIA. F	P	3,58	3,46	3,64	3,74	3,71	3,9	BELUM	BEKERJA	129	TERLAMBAT
C1255201008	WISNU DARMAJI	L	3,00	3,00	3,09	3,00	3,52	3,55	BELUM	BELUM	129	TEPAT WAKTU
C1255201009	SILVIA WIDYA.O	P	3,05	2,68	2,7	2,71	3,19	3,68	MENIKAH	BEKERJA	125	TERLAMBAT
C1255201010	DEFRI ZULHANSYAH	L	3,37	3,64	3,5	3,57	3,79	3,64	BELUM	BEKERJA	129	TEPAT WAKTU
C1255201011	DIKY KRISTOPAN	L	2,26	2,94	1,63	1,18	2,65	2,00	BELUM	BELUM	105	TERLAMBAT
C1255201012	SELVITA	P	3,21	2,77	3,00	2,71	3,52	3,54	BELUM	BELUM	127	TERLAMBAT
C1255201013	TRI KURNIA. N	P	3,53	2,83	3,4	3,3	3,74	3,4	BELUM	BELUM	129	TERLAMBAT
C1255201014	DWI RAHAYU. A	P	3,42	3,27	3,45	3,39	3,74	3,64	BELUM	BELUM	131	TERLAMBAT
C1255201015	MARIA FRANSISKA	P	2,74	2,95	2,8	3,14	3,26	3,5	BELUM	BELUM	125	TERLAMBAT
C1255201016	NORHIKMAH	P	3,21	2,77	3,2	3,13	3,26	3,32	BELUM	BELUM	129	TERLAMBAT
C1255201018	PUTRI SARI RAHMAN	P	3,32	3,23	3,36	3,43	3,74	4	BELUM	BELUM	131	TERLAMBAT
C1255201019	BRYAN ROFSON	L	2,37	2,94	2,6	2,71	3,38	3,5	BELUM	BELUM	121	TERLAMBAT
C1255201020	APRIADY	L	3	2,77	2,6	2,5	3,52	3,5	BELUM	BELUM	124	TEPAT WAKTU
C1255201021	SANIATI	P	3,32	3,14	3,18	3,48	3,65	3	MENIKAH	BELUM	131	TERLAMBAT
C1255201022	SAPTO PAMUNGKAS	L	3,22	3,52	3,5	3,94	1,2	0	BELUM	BEKERJA	146	TERLAMBAT
C1255201023	ERWIN SETIAWAN	L	0	0	0	0	0	0	BELUM	BELUM	29	DROP OUT
C1255201024	RONIE	L	2,16	2,44	2,25	2,41	3	2,86	BELUM	BELUM	109	TERLAMBAT
C1255201025	ARDI	L	2,79	2,78	2,4	2,06	2,05	2,17	BELUM	BELUM	112	DROP OUT
C1255201026	MUHAMMAD IRSYAD. P	L	2,84	3,4	2,91	2,86	3,62	3,67	BELUM	BELUM	127	TEPAT WAKTU

Lampiran 2. Data Uji

NIM	NAMA	JK	IPS_1	IPS_2	IPS_3	IPS_4	IPS_5	IPS_6	S_NIKAH	S_KERJA	J_SKS	SK
C1455201001	MAHANIS SEPTIANI	P	3,1	3	3,67	3,5	3,5	4	BELUM	BELUM	136	TEPAT WAKTU
C1455201004	TRI YOENATAN A BAYAN	L	0	0	3,09	3	0	0	BELUM	BELUM	155	TERLAMBAT
C1455201007	DIGDOYO PRADANA NURKHOLIS	L	3,71	3,17	3,73	3,71	3,42	4	BELUM	BELUM	128	TEPAT WAKTU
C1455201008	KLARA DWI LIGWINA	P	2,86	2,86	3,39	3,75	3,54	3,64	BELUM	BEKERJA	133	TEPAT WAKTU
C1455201009	MARCOPRIUS EGI SANJAYA	L	0,95	0	0	2	0	0	BELUM	BELUM	45	DROP OUT
C1455201010	RESKI ADVENTUS PRANATA	L	2	0,74	0	0	0	0	BELUM	BELUM	38	DROP OUT
C1455201015	AHMAT SEPTIYUNUS	L	3,38	3	3,12	3,5	3,77	3,86	BELUM	BELUM	126	TEPAT WAKTU
C1455201016	MUHAMMAD NUREFFENDI	L	3,71	3,09	3,12	3,42	3,68	3,64	BELUM	BELUM	126	TEPAT WAKTU
C1455201017	FARIKA SUCI NURHIKMAH	P	2,57	2,38	0,44	0	0	0	BELUM	BELUM	60	DROP OUT
C1455201021	PEBRIANSYAH	L	2,95	2,86	2,7	3,27	3,09	3,57	BELUM	BELUM	127	TEPAT WAKTU
C1455201022	DWI ARIANTO AKBAR	L	3,38	3,13	3,55	3,67	3,58	3,57	BELUM	BELUM	128	TERLAMBAT
C1455201026	JUMANTO	L	3,67	3,48	4	3,92	3,58	3,82	BELUM	BEKERJA	131	TEPAT WAKTU
C1455201027	INA MARYANA	P	2,9	2,95	3,9	3,54	3,67	3,42	BELUM	BEKERJA	129	TERLAMBAT
C1455201029	ERNIANDY SUGAWA	L	2,9	2,67	3,3	3,36	3,24	3,52	BELUM	BELUM	129	TEPAT WAKTU
C1455201034	KOMARIAH	P	2,95	3,1	3,28	3,58	3,54	3,61	BELUM	BELUM	128	TEPAT WAKTU
C1455201036	KHUSNUL DELIMAWATI	P	3,38	3,13	3	3,3	3,32	3,7	MENIKAH	BELUM	126	TERLAMBAT
C1455201037	GLORITA	P	2,48	2,63	3,6	3,33	3,55	3,43	BELUM	BELUM	129	TEPAT WAKTU
C1455201038	AHMAD MARUKI	L	2,9	2,62	2,7	3,18	3,18	3	BELUM	BELUM	124	TEPAT WAKTU
C1455201039	FITRIA NINGSIH	P	2,76	2,76	3,4	3,45	3,64	3,48	BELUM	BELUM	127	TEPAT WAKTU
C1455201041	PENIE	P	0,38	0	0	2	0	0	BELUM	BELUM	45	DROP OUT
C1455201042	SUCI FEBRINA	P	3,29	3,04	3,12	3,71	3,46	3,79	BELUM	BELUM	128	TEPAT WAKTU
C1455201044	MARIEYANTI LOE	P	2,57	2,67	3,4	2,73	2,05	2,3	BELUM	BEKERJA	124	TEPAT WAKTU
C1455201046	MUHAMMAD ABDULNASER	L	2,86	0,16	0	0	0	0	BELUM	BELUM	42	DROP OUT

C. Pengujian Menu *Testing 2*

Pengujian ini ketika pengguna ingin melakukan prediksi kelulusan Mahasiswa tanpa *file excel*. Tabel 3 merupakan hasil pengujian menu *testing 2*.

Tabel 3. Hasil Pengujian menu *Testing 2*

Hasil Pengujian			
Data Masukkan	Harapan	Pengamatan	Kesimpulan
Menginput <i>file tree</i>	Sistem akan melakukan pencarian <i>file tree</i>	Pengguna harus mengklik tombol <i>load tree</i>	Sesuai
Menginput data atribut	Sistem akan membaca data yang telah diinput, jika data tidak terisi dengan benar maka sistem akan menampilkan kotak dialog untuk mengisi data dengan benar	Pengguna harus menginput data secara manual	Sesuai, Namun sistem menanggapi jika di masukan angka lebih dari 4
Melakukan prediksi	Sistem akan menampilkan hasil prediksi	Pengguna harus mengklik tombol <i>prediction</i>	Sesuai
Mengatur ulang proses pelatihan data	Sistem akan mengosongkan data <i>inputan</i> dan <i>output</i>	Pengguna harus mengklik tombol <i>reset</i>	Sesuai

Data masukkan	Harapan	Pengamatan	Kesimpulan
Mengatur ulang proses pelatihan data	Sistem akan mengosongkan tabel, hasil prediksi dan hasil perhitungan akurasi algoritma	Pengguna harus mengklik tombol <i>reset</i>	Sesuai

B. Pengujian Menu *Testing 1*

Pengujian ini ketika pengguna melakukan *testing data*. Tabel 2 merupakan hasil pengujian menu *Testing 1*.

Tabel 2. Hasil Pengujian Menu *Testing 1*

Hasil Pengujian			
Data Masukkan	Harapan	Pengamatan	Kesimpulan
Menginput <i>file tree</i>	Sistem akan melakukan pencarian <i>file tree</i>	Pengguna harus mengklik tombol <i>load tree</i>	Sesuai
Menginput <i>file excel</i>	Sistem akan melakukan pencarian file dan membaca data yang ada di <i>file excel</i> lalu menampilkannya ditabel sistem	Pengguna harus mengklik tombol <i>load data</i>	Sesuai
Melakukan prediksi	Sistem akan menampilkan hasil prediksi dan hasil perhitungan akurasi algoritma.	Pengguna harus mengklik tombol <i>prediction</i>	Sesuai
Mengatur ulang proses pelatihan data	Sistem akan mengosongkan tabel, hasil prediksi dan hasil perhitungan akurasi algoritma	Pengguna harus mengklik tombol <i>reset</i>	Sesuai

BERITA ACARA BLACK BOX

Pada hari ini, Rabu tanggal 24 bulan Juli tahun 2019 telah dilaksanakan uji coba program dalam penulisan tugas akhir:

Nama : TETI SUMARNI

NIM : C1555201023

Judul : IMPLEMATASI ALGORITMA C4.5 DALAM PREDIKSI
KELULUSAN MAHASISWA STMIK PALANGKA RAYA

Poin-poin yang dieskperimen adalah:

A. Pengujian Menu *Training*

Pengujian ini ketika pengguna melakukan *training* data. Tabel 1 merupakan hasil pengujian menu *training*.

Tabel 1. Hasil Pengujian Menu *Training*

Hasil Pengujian			
Data masukkan	Harapan	Pengamatan	Kesimpulan
Menginput file excel	Sistem akan melakukan pencarian file dan membaca data yang ada di file excel lalu menampilkannya di tabel sistem	Pengguna harus mengklik tombol <i>load</i> data	Sesuai
Melakukan pelatihan data	Sistem akan menampilkan hasil prediksi, pohon keputusan dan hasil perhitungan akurasi algoritma.	Pengguna harus mengklik tombol <i>prediction</i>	Sesuai
Menyimpan pohon keputusan	Sistem dapat menyimpan file pohon keputusan	Pengguna harus mengklik tombol <i>save tree</i>	Sesuai

Keterangan :

*) Berikan **Sesuai** jika aplikasi dapat dijalankan dan **tidak sesuai** jika aplikasi tidak bisa dijalankan.

Yang Menguji



ILIA SAKHARIA

Palangka Raya, 24 Juli.....2019

Mahasiswa yang di uji,



TETI SUMARNI

Tabel Z Distribusi Normal

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.5	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170