

**PENERAPAN ALGORITMA C4.5UNTUK KLASIFIKASI  
MINAT BACA MAHASISWA  
(Studi kasus, Universitas Muhammadiyah Gorontalo)**

**Oleh**

**AMRAN PONTOH**

**T3114047**

**SKRIPSI**

Untuk Memenuhi Salah Satu Syarat Ujian

Guna Memperoleh Gelar Sarjana



**PROGRAM SARJANA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS ICHSAN GORONTALO  
GORONTALO  
2021**

## **PENGESAHAN SKRIPSI**

### **PENERAPAN ALGORITMA C4.5 UNTUK KLASIFIKASI MINAT BACA MAHASISWA**

Oleh

**AMRAN PONTOH**

**T3114047**

## **SKRIPSI**

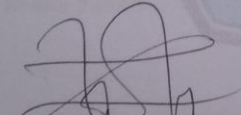
Untuk Memenuhi salah satu syarat ujian

Guna memperoleh gelar sarjana

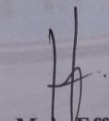
Program Studi Teknik Informatika dan telah disetujui oleh Tim Pembimbing

Gorontalo, 6 Juni 2021

**Pembimbing I**

  
Zohrahahaty, M.Kom  
NIDN. 0912117702

**Pembimbing II**

  
Moh. Effendi Lasulika, M.Kom  
NIDN. 0929048902

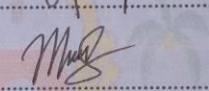
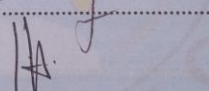
**PERSETUJUAN SKRIPSI**  
**PENERAPAN ALGORITMA C4.5 UNTUK KLASIFIKASI**  
**MINAT BACA MAHASISWA**

Oleh

AMRAN PONTOH

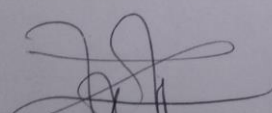
T3114047

Diperiksa Oleh Panitia Ujian Strata Satu (S1)  
Universitas Ichsan Gorontalo  
Gorontalo, 14 Juni 2021

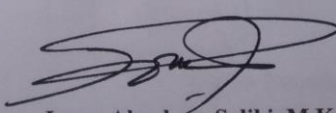
1. Ketua  
Irvan Abraham Salihi, M.Kom : 
2. Anggota  
Haditsah Annur, M.Kom : 
3. Anggota  
Muis Nanja, M.Kom : 
4. Anggota  
Zohrahayaty, M.Kom : 
5. Anggota  
Moh. Effendi Lasulika, M.Kom : 

Mengetahui

**Dekan Fakultas Ilmu Komputer**

  
Zohrahayaty, M.Kom  
NIDN. 0912117702

**Ketua Program Studi**

  
Irvan Abraham Salihi, M.Kom  
NIDN. 0928028101

## PERNYATAAN SKRIPSI

Denngan ini saya menyatakan:

- Karya tulis (Skripsi) saya ini adalah asli dan belum pernah diajukan untuk mendapatkan gelar akademik (Sarjana) Baik di universitas ichsan gorontalo maupun di perguruan tinggi lainnya.
- Karya tulis (Skripsi) saya ini adalah murni gagasan, rumusan, dan penelitian saya sendiri tanpa bantuan pihak lain kecuali arahan dari pembimbing.
- Dalam karya tulis (skripsi) saya ini tidak terdapat karya atau pendapat yang telah dipublikasi orang lain, kecuali secara tertulis dicantumkan pula dalam daftar pustaka.
- Pernyataan ini saya buat dengan sesungguhnya dan apa bila kemudian hari terdapat penyimpangan dan ketidak benaran dalam pernyataan ini maka saya bersedia menerima sangsi akademik berupa pencabutan gelar yang telah diperoleh karena karya tulis ini, serta sangsi lainnya sesuai dengan norma-norma yang berlaku di universitas ichsan gorontalo.

Gorontalo, 14 Juni 2021

Yang Membuat Pernyataan,

  
**Amran Pontoh**

## KATA PENGANTAR

*Assalamu'alaikum Wr.Wb*

Puji syukur penulis panjatkan kehadiran Allah SWT, karena atas berkat dan rahmat-Nya penulis dapat menyelesaikan Skripsi ini dengan judul, *Penerapan Algoritma C.45 Untuk Klasifikasi Minat Baca Mahasiswa*, sesuai dengan Apa yang direncanakan Penelitian ini dibuat untuk memenuhi salah satu syarat untuk ujian skripsi. Penulis menyadari tanpa bantuan Dari bimbingan berbagai pihak, Penelitian ini tidak dapat penulis diselesaikan. Oleh karena itu penulis menyampaikan Banyak terima kasih kepada :

1. Ibu Dra. Hj. Juriko Abdussamad, M.Si, Selaku Ketua Yayasan Pengembangan Ilmu Pengetahuan dan Teknologi (YPIPT) Ichsan Gorontalo.
2. Bapak Dr. Abdul Gaffar La Tjokke, M.Si., selaku Rektor Universitas Ichsan Gorontalo
3. Ibu Zohrahayaty, M.Kom., selaku Dekan Fakultas Ilmu Komputer sekaligus menjadi Pembimbing Utama, yang telah membimbing penulis selama mengerjakan Usulan Penelitian ini.
4. Bapak Sudirman S. Panna ,M.Kom., selaku Wakil Dekan I Bidang Akademik
5. Ibu Irma Surya Kumala, M.Kom., selaku Wakil Dekan II Bidang Administrasi Umum dan Keuangan
6. Bapak Sudirman Melangi, M.Kom, selaku Wakil Dekan III Bidang Kemahasiswaan
7. Bapak Irvan Abraham Salihi, M.Kom., selaku Ketua Program Studi Teknik Informatika Fakultas Ilmu Komputer
8. Bapak Moh. Effendi Lasulika, M.Kom., selaku Pembimbing Pendamping, yang telah membimbing penulis selama mengerjakan Skripsi ini.
9. Bapak dan Ibu Dosen yang telah mendidik dan membimbing penulis dalam mengerjakan Skripsi ini.
10. Seluruh keluarga, yang telah memberikan dukung moril dan doa kepada penulis

Semoga Allah, SWT melimpahkan balasan atas jasa-jasa mereka dan kepada penulis. Selanjutnya, penulis menyadari sepenuhnya bahwa apa yang telah dicapai ini masih jauh dari kesempurnaan dan masih banyak terdapat kekurangan. Oleh karena itu, penulis sangat mengharapkan adanya kritik dan saran yang konstruktif.

Akhirnya penulis berharap semoga hasil yang telah dicapai ini dapat mendukung program pemerintah dalam mencerdaskan kehidupan bangsa serta bermanfaat bagi pengembangan ilmu pengetahuan dan teknologi khususnya di bidang ilmu komputer maupun bermanfaat bagi masyarakat.

Akhir Kata, semoga bantuan dan amal kebaikan kepada penulis mendapatkan imbalan pahala dari keridahan dari ALLAH SWT,  
*Wassalamu'alaikum Wr.Wb*

Gorontalo, 31 Mei 2021

**Penulis**

## **ABSTRACT**

### **AMRAN PONTOH. T3114047. THE APPLICATION OF ALGORITHM C4.5 TO CLASSIFY STUDENT'S READING INTEREST**

*The reading interest in Gorontalo needs to be enhanced so that the education quality will also be improved. Consequently, libraries should make policies to provide supporting facilities. The problem that occasionally comes about is that the student's reading interest at Universitas Muhammadiyah Gorontalo is considered as relatively low. This study aims to identify the model C4.5 which has a high accuracy to classify the students' reading interest and to find a reliable and applicable C4.5-based classification system on students' reading interest. The results of this study indicate that the result of performance and effectiveness of the classification system using C4.5 can be identified through calculating the accuracy by 76.62%. Also, the implementation of C4.5 method in the classification system on students reading ability is found to be applicable based on the calculation  $CC = 4$  in white box test.*



**Keywords:** *reading interest, C4.5, classification*

## ABSTRAK

**AMRAN PONTOH. T3114047. JUDUL SKRIPSI PENERAPAN ALGORITMA C4.5 UNTUK KLASIFIKASI MINAT BACA MAHASISWA.**

Minat baca di gorontalo perlu di tingkatkan sehingga kualitas pendidikan pun meningkat. Untuk itu pihak perpustakaan dapat mengambil kebijakan untuk menyediakan fasilitas-fasilitas pendukung. Masalah yang sering terjadi adalah Minat baca mahasiswa Universitas Muhammadiyah Gorontalo masih tergolong rendah. Tujuan dari penelitian ini untuk Memperoleh Model C 4.5 yang akurasi tinggi untuk klasifikasi minat baca mahasiswa dan Untuk Memperoleh Sistem Klasifikasi Minat Baca Mahasiswa Berbasis C 4.5 yang handal sehingga dapat diimplementasikan. Hasil dari penelitian ini adalah dapat diketahui hasil kinerja dan keefektifan dalam sistem klasifikasi dengan menggunakan metode C4.5, dengan mengukur akurasi sebesar 76.62%. dan implementasi metode C4.5 dalam sistem klasifikasi minat baca mahasiswa dapat di implementasikan berdasarkan perhitungan  $CC = 4$  pada pengujian *white box*.



Kata Kunci: Minat Baca, C4.5, Klasifikasi



## DAFTAR ISI

HALAMAN JUDUL .....	i
HALAMAN PENGESAHAN SKRIPSI.....	ii
HALAMAN PERSETUJUAN SKRIPSI .....	iii
PERNYATAAN SKRIPSI .....	iv
KATA PENGANTAR .....	v
ABSTRACT ENGLISH/INDONESIA.....	vi
DAFTAR ISI .....	vii
DAFTAR GAMBAR .....	viii
DAFTAR TABEL .....	ix
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Identifikasi Masalah .....	3
1.3 Rumusan Masalah .....	3
1.4 Tujuan Penelitian .....	4
1.5 Manfaat Penelitian .....	4
BAB II TINJAUAN PUSTAKA .....	6
2.1 Tinjauan Studi .....	6
2.2 Tinjauan Teori.....	8
2.2.1 Minat Baca .....	8
2.2.2 Kecerdasan Buatan.....	9

2.2.3 Data Mining .....	12
2.2.4 Algoritma C4.5 ..	15
2.2.5 Confusion Matrix .....	17
2.2.6 Siklus Hidup Pengembangan Sistem .....	18
2.2.7 Rekayasa Perangkat Lunak Berbasis Objek .....	19
2.2.8 Desain sistem .....	21
2.2.9 Seleksi Sistem .....	28
2.2.10 Perawatan Sistem .....	28
2.2.11 Teknik Pengujian Sistem .....	29
2.2.12 Pengujian Sistem White Box .....	30
2.2.13 Pengujian Sistem Black Box .....	33
2.2.14 Perangkat Lunak Pendukung .....	35
2.3 Kerangka Pikir .....	36
BAB III METODE PENELITIAN .....	37
3.1 Objek Penelitian ... ..	37
3.2 Metode Penelitian . .....	37
3.2.1 Pengumpulan Data .....	37
3.2.2 Pra Pengolahan .....	39
3.2.3 Validasi Prediksi .....	39
3.2.4 Evaluasi Model .....	39
3.3 Pengembangan Sistem .....	39
3.3.1 Tahap Desain.. ..	40
3.3.2 Tahap Produksi Atau Pembuatan .....	41

3.3.3 Tahap Pengujian .....	42
3.3.4 Pengujian Sistem .....	43
BAB IV HASIL PENELITIAN.....	44
4.1 Hasil Pengumpulan Data .....	44
4.2 Arsitektur Algoritma C4.5 .....	45
4.3 Penerapan Metode C4.5.....	46
4.4 Sistem Yang Diusulkan .....	50
4.5 Diagram Konteks.....	51
4.6 Diagram Berjenjang.....	51
4.7 Diagram Arus Data .....	52
4.7.1 DAD Level 0 .....	52
4.7.2 DAD Level 1 Proses 1 .....	53
4.7.3 DAD Level 1 Proses 2 .....	53
4.7.4 DAD Level 1 Proses 3 .....	54
4.8 Kamus Data.....	54
4.8.1 Kamus Data User .....	54
4.8.2 Kamus Data Latih .....	55
4.8.3 Kamus Data Uji .....	56
4.8.4 Kamus Data Gain .....	57
4.8.5 Kamus Data Keputusan .....	57
4.8.6 Kamus Data Hasil Klasifikasi .....	58
4.9. Arsitektur Klasifikasi Peminat Baca .....	59
4.10. Interface Desain Mekanisme User.....	59

4.11. Interface Desain .....	59
4.12 Mekanisme Input .....	60
1. Input Login Admin .....	60
2. Input Data Latih .....	60
3. Input Data Uji .....	61
4.13. Interface Ouput .....	61
4.14. Data Desain .... ..	62
4.15.Pengujian White Box .....	62
4.16.Flowchart Pengujian White Box .....	63
4.17.Flowgraph .....	64
4.18. Perhitungan Pengujian White Box .....	65
4.19. Pengujian Basis Path.....	65
4.20. Pengujian Black Box .....	66
BAB V HASIL DAN PEMBAHASAN.....	67
5.1 Pembahasan Model .....	67
1. Tampilan Pembuatan Model .....	67
2. Tampilan Hasil Akurasi .....	67
5.2. Pembuatan Sistem .....	68
1. Tampilan Login Admin .....	68
2. Tampilan Halaman Admin .....	68
3. Tampilan Halaman Menu Data User .....	69
4. Tampilan Data Latih .....	69
5. Tampilan Hasil Klasifikasi .....	70

BAB VI KESIMPULAN DAN SARAN .....	71
6.1. Kesimpulan .....	71
6.2. Tampilan Data Latih .....	71
DAFTAR PUSTAKA	

## DAFTAR GAMBAR

	Halaman
1.1 Jumlah Pengunjung Perpustakaan UMGO .....	1
2.1 Akar Ilmu Data Mining .....	14
2.2 Siklus Hidup Pengembangan Sistem .....	19
2.3 SDLC Berorientasi Objek Dengan Pendekatan Waterfall .....	21
2.4 Contoh Bagan Alir White Box .....	28
2.5 Contoh Grafik Alir White Box .....	29
2.6 Kerangka Pikir .....	36
3.3 Sistem Yang Di usulkan .....	37
4.1. Arsitektur Algoritma C4.5 .....	45
4.2 Root Dari Hasil Gain .....	49
4.3 Sistem Yang Diusulkan .....	50
4.4 Diagram Konteks .....	51
4.5 Diagram Berjenjang .....	51
4.6 DAD Level 0.....	52
4.7 DAD Level 1 Proses 1.....	53
4.8 DAD Level 1 Proses 2.....	53
4.9 DAD Level 1 Proses 3.....	54
4.10 Interface Desain .....	59
4.11 Interface Login Admin .....	60
4.12 Interface Data Latih .....	60
4.13 Interface Data Uji .....	61

4.14 InterfaceOuput .....	61
4.15 Flowgraph.....	63
4.16 Flowgraph.....	64
5.1 Hasil Pemodelan .....	67
5.2 Akurasi .....	67
5.3 Tampilan Login .....	68
5.4 Tampilan Halaman Admin .....	68
5.5 Tampilan Halaman User .....	69
5.6 Tampilan Halaman Data Lath .....	69
5.7 Tampilan Halaman Klasifikasi .....	70

## DAFTAR TABEL

	Halaman
2.2 Confusion Matrix .....	17
2.3 Daftar Simbol Bagan Alir Dokumen.....	23
2.4 Daftar Simbol Bagan Alir Dokumen.....	24
2.5 Perangkat Lunak Pendukung .....	26
3.1 Atribut Data Prediksi Minat Baca .....	35
3.2 Data Klasifikasi Minat Baca .....	35
4.1. Data Penelitian .....	44
4.2. Data Penerapan .....	46
4.3. Perhitungan Gain .....	46
4.4. Kamus Data User.....	54
4.5 Kamus Data Latih.....	55
4.6 Kamus Data Uji .....	56
4.7 Kamus Data Gain .....	57
4.8 Kamus Data Keputusan .....	57
4.9 Kamus Data Hasil Klasifikasi .....	58
4.10 Mekanisme User .....	59
4.11 Pengujian Path .....	65
4.12 Hasil Pengujian Black Box .....	66



## **DAFTAR LAMPIRAN**

- 1. Lampiran Script(Program)**
- 2. Surat Rekomendasi Telah Melakukan Penelitian**
- 3. Surat Keterangan Bebas Plagiasi**
- 4. Hasil Cek Turnitin**

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Hingga saat ini minat baca masyarakat Indonesia masih tergolong rendah. Orang-orang lebih cenderung menonton ketimbang membaca [1]. Untuk itu minat baca perlu di tingkatkan sehingga kualitas pendidikanpun meningkat. Untuk itu pihak perpustakaan dapat mengambil kebijakan untuk menyediakan fasilitas-fasilitas pendukung.



**Gambar 1.1: Grafik Jumlah Pengunjung Perpustakaan UMGO**

Sumber: (Perpustakaan Universitas Muhammadiyah Gorontalo)

Berdasarkan data grafik yang diperoleh dari Perpustakaan Universitas Muhammadiyah Gorontalo diketahui banyaknya pengunjung ke perpustakaan dari tahun ke tahun meningkat. Akan tetapi banyak sedikitnya jumlah kunjungan belum dapat menentukan minat baca dari mahasiswa. Oleh karena itu, klasifikasi minat baca perlu dilakukan dengan akurat agar dapat mendukung pengambilan keputusan terkait perpustakaan kampus.

Kemajuan teknologi informasi menyebabkan kemudahan bagi setiap orang memperoleh data dengan mudah bahkan cenderung berlebihan. Data yang sedemikian besar tentunya memiliki informasi yang tersembunyi di dalamnya, namun kemampuan manusia terbatas dalam menganalisis data atau menggali pengetahuan dari data tersebut. Pengetahuan tersebut tentunya sangat berguna untuk mendukung pengambilan kebijakan atau keputusan. Selain itu, kemampuan komputasi yang semakin canggih dan terjangkau, serta persaingan bisnis yang semakin kompetitif merupakan faktor-faktor lainnya mengapa Data Mining semakin berperan dalam mendukung pengambilan keputusan [2].

Beberapa metode *machine learning* yang dapat minat baca mahasiswa, antara lain: (1) Artificial Neural Network (ANN); (2) Support Vector Machine (SVM); (3) K-Nearest Neighbor (K-NN); (4) C4.5. Di antara metode-metode tersebut, SVM berada sekelas dengan ANN yang merupakan metode yang sering digunakan untuk prediksi, keduanya merupakan metode *supervised learning*. ANN memiliki kelebihan dalam memprediksi data yang berdimensi besar dan melakukan pelatihan terhadap seluruh data. Sementara SVM tidak melakukan pelatihan terhadap seluruh data sehingga dapat mempercepat waktu komputasi

dan memiliki kelebihan pada masalah yang hanya 2 *class*, namun lemah pada dimensi data yang besar, dan tidak cocok untuk multi *classification*[3]. dengan demikian SVM tidak dapat digunakan karena data prediksi minat baca merupakan masalah *multi classification*.

Dengan demikian, penelitian ini akan menggunakan metode C4.5 untuk klasifikasi minat baca mahasiswa karena sesuai karakter data minat baca mahasiswa yang seluruh variabelnya bertipe diskret, yaitu: (1) Jenis Kelamin; (2) Jenis Bacaan; (3) Tujuan Membaca; (4) Lingkungan Baca, kecuali variabel (5) Umur saja yang bertipe numerik. Sedangkan variabel outputnya, yaitu: Minat Baca yang terdiri dari label: (1) Tinggi; (2) Sedang; (3) Rendah.

Pemodelan C 4.5 untuk klasifikasi minat baca mahasiswa dilakukan menggunakan alat bantu PHP, MySQL. Untuk menguji kinerja sistemnya, digunakan metode *White Box Testing* dan *Black Box Testing*.

Berdasarkan berbagai pemaparan di atas, maka peneliti tertarik untuk melakukan penelitian dengan judul: **“Penerapan Algoritma C4.5 Untuk Klasifikasi Minat Baca Mahasiswa.”** Diharapkan penelitian ini dapat memberikan kontribusi berupa Model C 4.5 yang akurasi paling tinggi untuk klasifikasi minat baca mahasiswa dan sistemnya yang handal sehingga dapat diimplementasikan.

## 1.2 Identifikasi Masalah

Berdasarkan latar belakang, dapat diidentifikasi masalah dari dua sisi, yaitu dari sisi masalah umumnya dan dari sisi ilmu komputer, yaitu:

1. Minat baca mahasiswa Universitas Muhammadiyah Gorontalo masih tergolong rendah.
2. Klasifikasi minat baca mahasiswa dapat diselesaikan dengan pendekatan klasifikasi Data Mining. Algoritma klasifikasi C 4.5 yang handal dalam mengolah data bertipe diskret namun dapat pula mengolah data bertipe numerik, Model C 4.5 Ini untuk klasifikasi minat baca mahasiswa yang telah diperoleh dapat digunakan untuk mengembangkan sistemnya.

### **1.3 Rumusan Masalah**

Berdasarkan identifikasi masalah, maka dapat dirumuskan masalah sebagai berikut:

1. Bagaimana hasil uji coba Algoritma C 4.5 yang akurasi tinggi untuk klasifikasi minat baca mahasiswa?
2. Bagaimana kinerja Sistem Klasifikasi Minat Baca Mahasiswa Berbasis C 4.5 yang dapat diimplementasikan?

### **1.4 Tujuan Penelitian**

Berdasarkan rumusan masalah, maka tujuan penelitian ini adalah:

1. Untuk Memperoleh Model C 4.5 yang akurasi tinggi untuk klasifikasi minat baca mahasiswa.
2. Untuk Memperoleh Sistem Klasifikasi Minat Baca Mahasiswa Berbasis C 4.5 yang handal sehingga dapat diimplementasikan.

### **1.5 Manfaat Penelitian**

Manfaat penelitian ini adalah sebagai berikut:

1. Manfaat Teoritis: memberikan masukan bagi perkembangan ilmu pengetahuan dan teknologi, khususnya pada bidang ilmu komputer, yaitu berupa Model C 4.5 yang akurat untuk klasifikasi minat baca mahasiswa dan sistemnya yang handal.
2. Manfaat Praktisi: sebagai sumbangan pemikiran, karya, bahan pertimbangan, atau solusi bagi pemerintah daerah guna mendukung pengambilan keputusan dalam rangka mengatur tata ruang yang ada di Perpustakaan.
3. Manfaat Peneliti: penelitian ini juga diharapkan dapat memberi masukan terhadap para peneliti yang akan mengadakan penelitian serupa atau mengembangkan penelitian ini.

## BAB II

### TINJAUAN PUSTAKA

#### 2.1 Tinjauan Studi

Terdapat beberapa penelitian yang terkait tentang klasifikasi minat baca dan penerapan Algoritma C.45, seperti di bawah ini:

Penelitian	Judul	Teori
David Hartanto Kamagi  dkk (2014). [4]	Implementasi Data Mining dengan Algoritma C4.5 untuk klasifikasi Tingkat Kelulusan Mahasiswa.	1. Dari hasil penelitian yang telah dilakukan algoritma C4.5 dapat diimplementasikan untuk klasifikasi tingkat kelulusan mahasiswa dengan empat kategori yaitu lulus cepat, lulus tepat, lulus terlambat dan <i>drop out</i> . Attribute yang paling berpengaruh dalam hasil adalah IPS semester enam hasil kelulusan mahasiswa dengan presentase 87.5% dari enam puluh data training dan empat puluh data testing.

<p>Eki Ruziqa Maris (2014). [5]</p>	<p>Analisis Kepuasan Pelanggan Menggunakan Algoritma C4.5.</p>	<p>2. Dari hasil penelitian yang telah dilakukan <i>Ratio data training</i> yang digunakan mempengaruhi nilai akurasi pada setiap percobaan. pada percobaan ke-1 nilai akurasi adalah 86% dengan <i>data training 70% data testing 30%</i>. Pada percobaan ke-2 nilai akurasi adalah 93% <i>data training 90% dan data testing 10%</i>. Dengan hasil ini pengujian yang terbaik yaitu dengan akurasi 93%.</p>
<p>Wisti Dwi Septiani (2014). [6]</p>	<p>Penerapan Algoritma C4.5 Untuk klasifikasi Penyakit Hepatitis</p>	<p>3. Dari hasil penelitian yang telah dilakukan pada data pasien penderita penyakit hepatitis bahwa Algoritma C4.5 menghasilkan akurasi 77,29% dan nilai AUC 0,846 yang termasuk dalam <i>Good Classification</i>. Dengan demikian dapat disimpulkan bahwa metode ini akurat dalam melakukan hasil untuk penyakit hepatitis.</p>



Fina Nasari (2014). [7]	Penerapan Algoritma C4.5 Dalam Pemilihan Bidang Peminatan Program Studi Sistem Informasi Di Stimik Potensi Utama Medan.	4. Dari hasil penelitian yang telah dilakukan berdasarkan perhitungan menggunakan algoritma C4.5 diperoleh faktor dominan seseorang memilih bidang peminatan adalah berdasarkan nilai JK dengan tingkat kecocokan data hingga 80.14%.
----------------------------	---	---

## 2.2 Tinjauan Teori

### 2.2.1 Minat Baca

Membaca merupakan kegiatan dan kemampuan khas manusia. Namun demikian kemampuan membaca manusia tidak terjadi secara otomatis karena didahului oleh aktivitas dan kebiasaan membaca yang merupakan wujud adanya minat baca. Minat baca menurut [8] adalah suatu sikap yang memperlihatkan perhatian dan kemampuan pada membaca yang membuat membaca menjadi menyenangkan, penuh motivasi kontinu, focus dan berarti. Motivasi yang rendah merupakan salah satu penyebab kurangnya minat baca. Untuk menumbuhkan minat baca sangat dibutuhkan suatu upaya yang keras ada pada setiap individu, sebab minat dapat memiliki karena dibangkitkan atau ditimbulkan, baik dari diri sendiri, orang lain maupun oleh lingkungan universitas.

Minat baca adalah suatu ketertarikan untuk dapat mengartikan atau menafsirkan media kata-kata dengan tujuan untuk memperoleh informasi yang

dibutuhkan. Dengan adanya minat baca dapat mendorong seseorang untuk giat memperluas pengetahuannya. Semakin tinggi minat baca pada diri seseorang semakin tinggi pula hasil belajar yang diterimanya, sehingga diharapkan dapat mencapai tujuan belajar yang optimal.

### 2.2.2 Kecerdasan Buatan

Sebagian kalangan menerjemahkan *Artificial Intelligence* sebagai kecerdasan buatan, kecerdasan artifisial, inteligensia artifisial, atau inteligensia buatan. Para ahli mendefinisikan AI secara berbeda-beda tergantung pada sudut pandang mereka masing-masing. Ada yang fokus pada logika berpikir manusia saja, tetapi ada juga yang mendefinisikan AI secara lebih luas pada tingkah laku manusia. Pada [RUS95], Stuart Russel dan Peter Norvig mengelompokkan definisi AI, yang diperoleh dari beberapa *textbook* berbeda, ke dalam empat kategori, yaitu:

#### 1. *Thinking humanly: the cognitive modeling approach*

Pendekatan ini dilakukan dengan dua cara sebagai berikut:

- (a) Melalui introspeksi: mencoba menangkap pemikiran-pemikiran kita sendiri pada saat kita berpikir. Tetapi, seorang psikolog berat mengatakan "*how do you know that you understand?*" Bagaimana anda tahu bahwa anda mengerti? Karena pada saat anda menyadari pemikiran anda, ternyata pemikiran tersebut sudah lewat dan digantikan kesadaran anda. Sehingga, definisi ini terkesan mengada-ada dan tidak mungkin dilakukan.
- (b) Melalui eksperimen-eksperimen psikolog.

## 2. *Acting humanly: the turing test approach*

Pada tahun 1950, Alan Turing merancang suatu ujian bagi komputer berintelijensia untuk menguji apakah komputer tersebut mampu mengelabui seorang manusia yang menginterogasinya melalui *teletype* (komunikasi berbasis teks jarak jauh). Jika *interrogator* tidak dapat membedakan yang diinterogasi adalah manusia atau komputer, maka komputer berintelijensia tersebut lolos dari *Turing test*. Komputer tersebut perlu memiliki kemampuan: *Natural Language Processing, Knowledge Representation, Automated Reasoning, Machine Learning, Computer Vision, Robotics*. *Turing test* sengaja menghindari interaksi fisik antara *interrogator* dan komputer karena simulasi fisik manusia tidak memerlukan intelijensia.

## 3. *Thinking rationally: the laws of thought approach*

Terdapat dua masalah dalam pendekatan ini, yaitu:

- (a) Tidak mudah untuk membuat pengetahuan informal dan menyatakan pengetahuan tersebut ke dalam *formal term* yang diperlukan oleh notasi logika, khususnya ketika pengetahuan tersebut memiliki kepastian kurang dari 100%.
- (b) Terdapat perbedaan besar antara dapat memecahkan masalah “dalam prinsip” dan memecahkannya “dalam dunia nyata”.

## 4. *Acting rationally: the rational agent approach*

Membuat inferensi yang logis merupakan bagian dari suatu *rational agent*. Hal ini disebabkan satu-satunya cara untuk melakukan aksi secara rasional adalah dengan menalar secara logis. Dengan menalar secara logis, maka bisa

didapatkan kesimpulan bahwa aksi yang diberikan akan mencapai tujuan atau tidak. Jika mencapai tujuan, maka *agent* dapat melakukan aksi berdasarkan kesimpulan tersebut.

*Thinking humanly* dan *acting humanly* adalah dua definisi dalam arti yang sangat luas. Sampai saat ini, pemikiran manusia yang diluar rasio, yakni reflex dan intuitif (berhubungan dengan perasaan), belum dapat ditirukan sepenuhnya oleh komputer. Dengan demikian, kedua definisi ini dirasa kurang tepat untuk saat ini. Jika kita menggunakan definisi ini, maka banyak produk komputasi cerdas saat ini yang tidak layak disebut sebagai produk AI. Definisi *thinking rationally* terasa lebih sempit dari pada *acting rationally*. Oleh karena itu, definisi AI yang paling tepat untuk saat ini adalah *acting rationally* dengan pendekatan *rational agent*. Hal ini berdasarkan pemikiran bahwa komputer bisa melakukan penalaran secara logis dan juga bisa melakukan aksi secara rasional berdasarkan hasil penalaran tersebut.

Para ilmuwan memiliki dua cara pandang yang berbeda tentang AI. Yang pertama memandang AI sebagai bidang ilmu yang hanya fokus pada proses berpikir. Sedangkan yang ke dua memandang AI sebagai bidang ilmu yang fokus pada tingkah laku. Cara pandang ke dua melihat AI secara lebih luas karena suatu tingkah laku pastilah didahului dengan proses berpikir. Saat ini, dengan semakin cepatnya perkembangan *hardware* dan *software*, berbagai produk AI telah berhasil dibangun dan digunakan dalam kehidupan sehari-hari. Dengan teknologi *hardware* yang performansinya semakin tinggi dan berukuran kecil serta didukung teknologi *software* yang semakin beragam dan kuat, produk-produk

berbasis AI semakin dekat dengan kehidupan manusia. Pada masa mendatang, AI ditantang untuk membuat suatu kecerdasan yang hamper menyamai kecerdasan manusia. Ray Kurzweil memprediksi bahwa hal itu akan mungkin terwujud melalui tahapan-tahapan yang dibuatnya secara bertahap hingga tahun 2099.[9]

### **2.2.3 Data Mining**

*Data mining* merupakan cabang ilmu yang termasuk masih baru tetapi telah menghasilkan keuntungan yang cukup besar saat ini. Manfaat yang dirasakan dengan penggunaan *Data Mining* sudah tidak diragukan lagi. Beberapa kasus telah banyak diselesaikan dengan menggunakan teknik data *mining* ini. Misalnya kasus peledakan gedung WTC oleh teroris ternyata telah terdeteksi oleh pihak Amerika Serikat, hanya saja sayangnya terlambat untuk diantisipasi.

*Data Mining* sudah sering dimanfaatkan untuk mendeteksi kejadian-kejadian yang ganjil seperti penyakit tertentu, transaksi yang mencurigakan, hingga mendeteksi telepon yang dilakukan oleh pihak yang bermaksud menipu, seperti penyalahgunaan kartu kredit. Pihak pasar swalayan pun saat ini sudah menggunakan teknik *Data Mining* untuk mengatur peletakan *stand* penjualan agar mempermudah pembeli dalam pergerakannya di dalamnya.

Saat ini sudah banyak vendor-vendor yang mengkhususkan diri dalam bidang *Data Mining* seperti SPSS, Microsoft, Oracle, dan lain-lain. Produknya pun beragam dari bidang ekonomi seperti *Bussines Intelligent* (BI), *Business Performance Management* (BPM), *Customer Relationship Management* (CRM), dan sebagainya hingga bidang kesehatan seperti *Health Informatics*, *e-Health*, dan

sebagainya. Tentu saja bidang yang lain juga telah lama menerapkan *Data Mining* seperti Telekomunikasi, Meteorologi, Biologi dan lain-lain.[10]

Data mining juga sering disebut sebagai kegiatan mengeksplorasi dan menganalisis data dalam jumlah yang besar untuk menemukan pattern dan rule yang bearti. Data mining digunakan untuk mencari informasi bisnis berharga dari basis data yang sangat besar yang dipakai untuk memprediksi tren dan sifat-sifat bisnis serta menemukan pola-pola yang tidak diketahui sebelumnya. Berdasarkan beberapa arti dan pengertian mengenai data mining, maka dapat diambil suatu kesimpulan mengenai data mining. Jadi data mining adalah proses pencarian informasi dan pengetahuan dari keseluruhan data yang tersembunyi, dengan menggunakan teknik yang sesuai dengan tempat yang digali sehingga nantinya akan menemukan suatu pola dari suatu data yang mungkin selama ini terabaikan dan dianggap sebagai suatu hal yang tak penting untuk kemudian diimplementasikan dalam bentuk grafik maupun diagram serta bentuk jenis lainnya yang mempunyai inti yang berisikan pengetahuan dan informasi yang berguna bagi penggunaannya.[11]

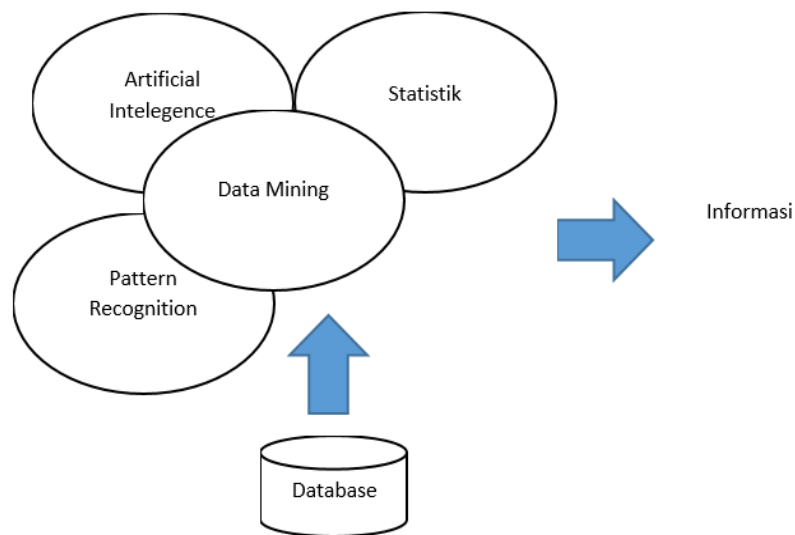
Nama data mining sebenarnya muali dikenal sejak tahun 1990, ketika pekerjaan pemanfaatan data menjadi sesuatu yang penting dalam berbagai bidang, mulai dari bidang akademik, bisnis, hingga medis. [12] Data mining dapat diterapkan pada berbagai bidang yang mempunyai sejumlah data, tetapi karena wilayah penelitian dengan sejarah yang belum lama, dan belum melewati masa ‘remaja’ maka data mining masih diperdebatkan posisi bidang pengetahuan yang memilikinya. Maka, Daryl Pregibon menyatakan bahwa “data mining adalah

campuran dari statistic, kecerdasan buatan, dan riset basis data” yang masih berkembang.[12]

Terlepas dari ‘remaja’-nya data mining, ternyata data mining diproyeksikan menjadi jutaan dolar di dunia industry pada tahun 2000, sedangkan pada saat yang sama, ternyata data mining dipandang sebelah mata oleh sejumlah peneliti sebagai *dirty word in statistic*. [12] Mereka adalah orang-orang yang tidak memandang data mining sebagai sesuatu yang menarik bagi mereka pada saat itu.

Munculnya data mining didasarkan pada jumlah data yang tersimpan dalam basis data semakin besar. Misalnya dalam sebuah supermarket, ada berapa transaksi pelanggan yang terjadi dalam sehari dan ada berapa juta data yang sudah tersimpan dalam sebulan. Dalam perusahaan, ada berapa juta data yang sudah tersimpan dari setiap kegiatan produksi untuk setiap produk yang dibuat dalam beberapa tahun. Contoh lain, jika anda mempunyai kartu kredit, mungkin anda sering menerima surat penawaran barang atau jasa. Jika bank mempunyai 1.000.000 nasabah dan biaya pengiriman surat pernasabah adalah 500 rupiah, maka biaya yang harus dikeluarkan bank adalah 500 juta rupiah, padahal nasabah yang mungkin benar-benar membeli hanya sekitar 15%. Akibatnya, ada pembuangan biaya sekitar 85% dari 500 juta atau sekitar 425 juta; sungguh sia-sia. Jika perusahaan dapat memanfaatkan data-data yang ada sehingga hanya nasabah yang berpotensi untuk membeli saja yang dikirim surat, maka biaya pengiriman tersebut dapat ditekan.

Yang menjadi pertanyaan untuk data-data dalam perusahaan yang semakin menggunung jumlahnya dari waktu ke waktu adalah mau diapakan data tersebut, apakah hanya untuk dibuat laporan akhir tahun kemudian dibuang? Apakah hanya akan dikubur dalam gudang data dan tidak diapa-apakan? Tentu sayang sekali jika data-data tersebut tidak dimanfaatkan untuk kepentingan perusahaan atau instansi-instansi yang berkepentingan.[2]



**Gambar2.1 Akar ilmu data mining**

Sumber (Prasetyo, 2014)

#### 2.2.4 Algoritma C 4.5

Algoritma C4.5 diperkenalkan oleh Quinlan (1996) sebagai versi perbaikan dari ID3. Dalam ID3, induksi decision tree hanya bias dilakukan pada fitur bertipe kategorikal (nominal atau ordinal), sedangkan tipe numerik (interval atau rasio) tidak dapat digunakan. Perbaikan yang membedakan algoritma C4.5 dari ID3 adalah dapat menangani fitur dengan tipe numerik, melakukan pemotongan (*pruning*) decision tree, dan penurunan (*deriving*) rule set. Algoritma



C4.5 juga menggunakan kriteria gain dalam menentukan fitur yang menjadi pemecah node pada pohon yang di induksi.

Secara umum Algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut:

- a. Pilih atribut sebagai akar
- b. Buat cabang untuk masing-masing nilai
- c. Bagi kasus dalam cabang
- d. Ulangi proses untuk 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:

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

Keterangan:

S = Himpunan kasus

A = Atribut

n = jumlah partisi atribut A

|S<sub>i</sub>| = jumlah kasus pada partisi ke i

|S| = jumlah kasus dalam S

Sebelum mendapatkan nilai gain adalah dengan mencari nilai entropi. Entropi digunakan untuk menentukan seberapa informatif sebuah masukan atribut untuk menghasilkan sebuah atribut. Rumus dasar dari entropi adalah sebagai berikut:

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

Keterangan:

S = Himpunan Kasus

n = Jumlah partisi S

$p_i$  = Proporsi dari  $S_i$  terhadap S.

### 2.2.5 Confusion Matrix

Confusion Matrix adalah suatu metode yang biasanya digunakan untuk melakukan perhitungan akurasi pada konsep data mining. Confusion matrix memberikan keputusan yang diperoleh penilaian performance klasifikasi berdasarkan objek dengan benar atau salah.[12] Confusion matrix berisi informasi actual (actual) dan prediksi (predicted) pada sistem klasifikasi.

**Tabel 2.2 Tabel Confusion Matrix**

	actual	Actual
	positive	Negative
predicted positive	<i>TP</i>	<i>FP</i>
predicted negative	<i>FN</i>	<i>TN</i>

Recall	=	$TP/(TP + FN)$
Precision	=	$TP/(TP + FP)$
True Positive Rate	=	$TP/(TP + FN)$
False Positive Rate	=	$FP/(FP + TN)$

Sumber: (gurunescu, 2011)

Keterangan:

- (a) True Positive (tp) = proporsi positif dalam data set yang dikalsifikasikan positif.
- (b) True Negative (tn) = proporsi negative dalam data set yang diklasifikasikan negative.
- (c) False Positive (fp) = proporsi negative dalam data set yang diklasifikasikan positive.
- (d) False Negative (fn) = proporsi negative dalam data set yang diklasifikasikan

#### **2.2.6 Siklus Hidup Pengembangan Sistem**

Pengembangan sistem informasi yang berbasis komputer dapat merupakan tugas kompleks yang membutuhkan banyak sumber daya dan dapat memakan waktu berbulan-bulan bahkan bertahun tahun untuk menyelesaikannya. Proses pengembangan sistem melewati beberapa tahapan dari mulai sistem itu direncanakan sampai dengan sistem tersebut diterapkan, dioperasikan dan dipelihara. Bila operasi sistem yang sudah dikembangkan masih timbul kembali permasalahan-permasalahan yang kritis serta tidak dapat diatasi dalam tahap pemeliharaan sistem, maka perlu dikembangkan kembali suatu sistem untuk mengatasinya dan proses ini kembali ke tahap yang pertama, yaitu tahap perencanaan sistem. Siklus ini disebut dengan siklus hidup suatu sistem (*systems development life cycle*). Daur atau siklus hidup dari pengembangan sistem merupakan suatu bentuk yang digunakan untuk menggambarkan tahapan utama dan langkah-langkah didalam tahapan tersebut dalam proses pengembangannya.

Tahapan utama siklus hidup pengembangan sistem dapat terdiri dari tahapan perencanaan sistem, analisis sistem, desain sistem, seleksi sistem, implementasi sistem dan perawatan sistem. Tahapan-tahapan seperti ini sebenarnya merupakan tahapan di dalam pengembangan sistem teknik.[13]



**Gambar 2.2 Siklus Pengembangan Sistem**

Sumber; jogiyanto HM,(2005:52)

### 2.2.7 Rekayasa Perangkat Lunak Berbasis Objek

*Software* (perangkat lunak) berhubungan dengan: (1) Perintah (program komputer) yang bila dieksekusi memberikan fungsi dan unjuk kerja seperti yang diinginkan; (2) Struktur data yang memungkinkan program memanipulasi informasi secara proporsional; dan (3) dokumen yang menggambarkan operasi dan kegunaan program. “*Software engineering* adalah suatu disiplin rekayasa (rancang-bangun) yang terkait dengan semua aspek produksi perangkat lunak (Pressman, Rekayasa Perangkat Lunak (Terjemahan), 2002)”. Sedangkan menurut IEEE (Institute of Electrical and Electronic Engineers):

“Software engineering adalah: (1) Aplikasi dari sebuah pendekatan yang bersifat kuantifiabel, disiplin, dan sistematis bagi pengembangan, operasi, dan pemeliharaan perangkat lunak; (2) Studi tentang pendekatan-pendekatan seperti pada (1). [14]

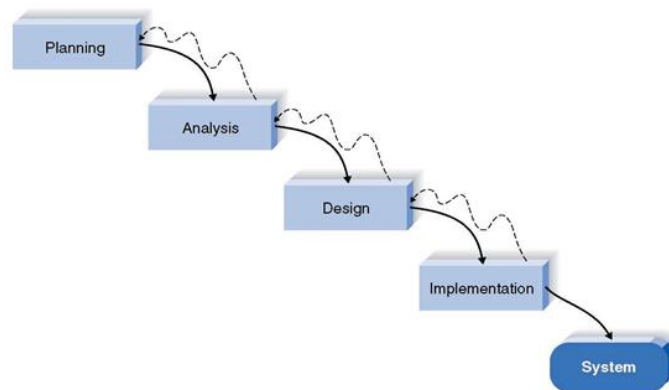
Dengan demikian bidang penelitian *software engineering* mengacu pada kedua hal tersebut.

Perbedaan utama paradigma berorientasi objek dibandingkan konvensional adalah pada penyatuan proses/fungsi dan data ke dalam bentuk yang terenkapsulasi, sedangkan paradigma konvensional memisahkan data dengan proses. Data merupakan sekumpulan atribut yang di-enkapsulasi (dikemas/dibungkus) bersama algoritma (operasi/metode/servis) untuk melakukan suatu proses berdasarkan pesan (stimulan/message/event) yang masuk ke objek/*class*. Pesan ini merupakan sarana *interface* antar-objek.[14]

Objek-objek merupakan realitas yang unik sehingga tidak efisien bila dideskripsikan satu persatu sementara mereka memiliki beberapa kesamaan. Misal kerbau dengan sapi adalah dua objek yang berbeda, tetapi keduanya adalah pemakan rumput. Oleh sebab itu diciptakan kelas yang memuat karakteristik yang sama dari kerbau dan sapi. Dengan demikian kelas ini dapat menginformasikan berbagai binatang pemakan rumput (tetapi tidak bisa menyebutkan binatang-binatang di kebun binatang). Kelas ini dengan mudah dikembangkan untuk merekrut objek/binatang-binatang lain pemakan rumput secara efisien tanpa mengulang-ulang deskripsi yang sama. Bahkan dapat dibuat hirarkhinya sebagai superkelas atau subkelas yang masing-masing memiliki warisan (*inheritance*) dari

kelas di atasnya. Sedangkan *Polymorphism* diperlukan guna memperluas sistem (software engineering) berbasis objek yang ada secara efisien, yakni dengan nama yang sama tetapi operasinya berbeda.[14]

*Sistem Development Life Cycle* (SDLC) berorientasi objek dengan pendekatan *waterfall* dapat ditunjukkan pada gambar berikut ini.



**Gambar 2.3 SDLC Berorientasi Objek dengan Pendekatan Waterfall**

Sumber: (Alan Dennis, 2010)

### 2.2.8 Desain Sistem

Dalam desain sistem, dibutuhkan alat bantu desain. Dalam tahapan ini, pengembang sistem bisa menentukan arsitektur sistemnya, merancang gambaran konseptual sistem, merancang database, perancangan interface, hingga membuat flowchart program. Salah satu alat bantu yang bisa digunakan dalam pembuatan sistem bantu keputusan adalah *Data Flow Diagram* (DFD). DFD adalah suatu model logika data atau proses yang dibuat untuk menggambarkan asal data dan tujuan data yang keluar dari sistem, tempat penyimpanan data, proses apa yang menghasilkan data tersebut, serta interaksi antara data yang tersimpan dan proses yang dikenakan pada data tersebut.

Menurut John Burch dan Gary Grudnitski, desain sistem dapat didefinisikan sebagai penggambaran, perencanaan, dan pembuatan sketsa atau pengaturan dari beberapa elemen yang terpisah ke dalam satu kesatuan yang utuh dan berfungsi. [13]

Tahap desain sistem mempunyai dua tujuan utama :

1. Untuk memenuhi kebutuhan kepada pemakai sistem.
2. Untuk memberikan gambaran yang jelas dan rancang bangun yang lengkap kepada pemrogram komputer dan ahli-ahli teknik lainnya.

Desain sistem dibagi dalam dua bagian, yaitu desain sistem secara umum (*general systems design*) dan desain sistem terinci (*detailed systems design*).

#### **1. Desain Sistem Secara Umum (*general systems design*)**


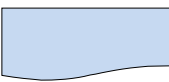

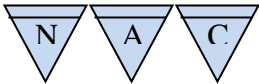



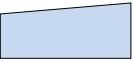

Pada tahap desain secara umum, komponen-komponen sistem informasi yang dirancang dengan tujuan dikomunikasikan kepada user bukan untuk pemrograman. Komponen sistem informasi yang di desain adalah model, output, input, database, teknologi, dan kontrol. [13]

##### **a. Desain Model Secara Umum**


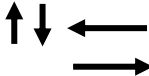

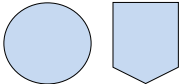
Analisis sistem dapat mendesain model dari sistem informasi yang di usulkan dalam bentuk *physical* sistem dan *logical* model. Bagan alir sistem merupakan alat yang tepat digunakan untuk menggambarkan *physical systems*, *logical model* dapat digambar dengan diagram arus data. [13]

Bagan alir sistem merupakan bagan yang menunjukkan arus pekerjaan secara keseluruhan dari sistem. Bagan alir sistem digambar dengan simbol-simbol sebagai berikut :

**Tabel 2.3**Daftar Simbol Bagan Alir Dokumen

No.	Nama Simbol	Simbol	Keterangan
1.	Terminal		Menunjukkan untuk memulai dan mengakhiri Suatu proses
2.	Dokumen		Menunjukkan dokumen input dan output baik itu proses manual, mekanik, atau computer
3.	Kegiatan Manual		Menunjukan pekerjaan manual
4.	Simpanan Offline		Menunjukkan file non-komputer yang diarsip urut angka ( <i>numerical</i> ), huruf ( <i>alphabetical</i> ), atau tanggal ( <i>chronological</i> )
5.	Proses		Menunjukkan kegiatan proses dari operasi program komputer
6.	Operasi Luar		Menunjukkan operasi yang dilakukan diluar operasi computer
7.	Hard Disk		Menunjukkan <i>input</i> dan <i>output</i> menggunakan <i>harddisk</i>
8.	Keyboard		Menunjukkan <i>input</i> yang menggunakan <i>on-line keyboard</i>
9.	Display		Menunjukkan <i>output</i> yang ditampilkan di monitor

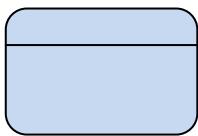



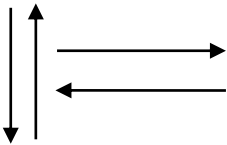

No.	Nama Simbol	Simbol	Keterangan
10	Hubungan Komunikasi		Menunjukkan proses transmisi data melalui channel komunikasi
11	Garis Alir		Menunjukkan arus dari proses
12	Penjelasan		Menunjukkan penjelasan dari suatu proses
13	Penghubung		Menunjukkan penghubung ke halaman yang masih sama atau ke halaman yang lain

(Sumber: Jogiyanto HM, 2005)

Untuk mempermudah penggambaran suatu sistem yang ada atau sistem yang baru yang akan dikembangkan secara logika dan tanpa memperhatikan lingkungan fisik data tersebut mengalir atau lingkungan fisik dimana data tersebut akan disimpan, maka digunakan Diagram Arus Data (DAD) atau *Data Flow Diagram* (DFD).

**Tabel 2.4**Daftar Simbol Diagram Alir Dokumen

No	Simbol	Keterangan
1.		Simbol Proses, Menunjukkan informasi dari masukan menjadi keluaran

No	Simbol	Keterangan
2.		Eksternal Entity, merupakan kesatuan dilingkungan luar system yang dapat berupa orang, organisasi atau system lain yang berada di lingkungan luarnya yang akan memberikan input seta menerima output dari system
3.		Aliran atau arus data, menggambarkan gerakan paket data atau informasi dari suatu bagian kebagian yang lain, dimana penyimpanan mewakili lokasi penyimpanan data
4.		Penyimpanan, digunakan untuk memodelkan kumpulan data atau paket data

(Sumber : Jogiyanto, 2005)

#### b. Desain Output Secara Umum

Output adalah produk dari sistem informasi yang dapat dilihat. Output terdiri dari macam-macam jenis seperti hasil di media kertas, dan hasil di media lunak. Disamping itu output dapat berupa hasil dari suatu proses yang akan digunakan oleh proses lain dan tersimpan di suatu media seperti tape, disk, atau kartu. Yang dimaksud dengan output pada tahap desain ini adalah output yang berupa tampilan di media kertas atau di layar video. [13]

#### c. Desain Input Secara Umum

Alat input dapat digolongkan ke dalam 2 golongan, yaitu alat input langsung (*online input device*) dan alat input tidak langsung (*offline input device*). Alat input langsung merupakan alat input yang langsung dihubungkan dengan CPU, sedangkan alat input tidak langsung adalah alat input yang tidak langsung dihubungkan dengan CPU. [13]

d. Desain Database Secara Umum

Basis data (database) adalah kumpulan dari data yang saling berhubungan satu dengan yang lainnya, tersimpan diluar komputer dan digunakan perangkat lunak tertentu untuk memanipulasinya. Sistem basis data adalah suatu sistem informasi yang mengintegrasikan kumpulan dari data yang saling berhubungan satu dengan yang lainnya dan membuatnya tersedia untuk beberapa aplikasi yang bermacam-macam di dalam suatu organisasi. [13]

**2. Desain Sistem Secara Rinci (*Detailed systems design*)**

a. Desain Output Terinci

Desain output terinci dimaksudkan untuk mengetahui bagaimana dan seperti apa bentuk output-output dari sistem yang baru. Desain output terinci terbagi atas dua, yaitu desain output berbentuk laporan di media kertas dan desain output dalam bentuk dialog di layar terminal. [13]

1. Desain output dalam bentuk laporan : dimaksudkan untuk menghasilkan output dalam bentuk laporan di media kertas. Bentuk laporan yang paling

banyak digunakan adalah dalam bentuk tabel dan berbentuk grafik atau bagan. [13]

2. Desain output dalam bentuk dialog layar terminal : merupakan rancang bangun dari percakapan antara pemakai sistem atau user dengan komputer. Percakapan ini dapat terdiri dari proses memasukkan data ke sistem, menampilkan output informasi kepada user, atau keduanya.

b. Desain Input Terinci

Masukan merupakan awal dimulainya proses informasi. Bahan mentah dari informasi adalah data yang terjadi dari transaksi-transaksi yang dilakukan oleh organisasi. Data hasil dari transaksi merupakan masukan untuk sistem informasi. Hasil dari sistem informasi tidak lepas dari data yang dimasukan. Desain input terinci dimulai dari desain dokumen dasar sebagai penangkap input yang pertama kali. Jika dokumen dasar tidak di desain dengan baik, kemungkinan input yang tercatat dapat salah bahkan kurang. [13]

Fungsi dokumen dasar dalam penanganan arus data :

1. Dapat menunjukkan macam dari data yang harus dikumpulkan dan ditangkap.
2. Dapat dicatat dengan jelas, konsisten, dan akurat.
3. Dapat mendorong lengkapnya data, disebabkan data yang dibutuhkan disebutkan satu persatu di dalam dokumen dasarnya.

c. Desain Database Terinci

Database merupakan kumpulan dari data yang saling berhubungan satu dengan yang lainnya, tersimpan di simpanan luar komputer dan digunakan perangkat lunak tertentu untuk memanipulasinya. Database merupakan salah satu komponen yang penting di dalam sistem informasi, karena berfungsi sebagai basis penyedia informasi bagi para pemakainya. Penerapan database dalam sistem informasi disebut *database system*. [13]

### **2.2.9 Seleksi Sistem**

Tahap ini merupakan tahap untuk memilih perangkat yang akan digunakan untuk sistem informasi. Pengetahuan dibutuhkan oleh pemilih sistem diantaranya adalah pengetahuan tentang siapa yang menyediakan teknologi ini, cara pemilikannya, dan sebagainya. Pemilihan sistem yang harus paham dengan teknik-teknik evaluasi untuk menyelesaikan sistem

### **2.2.10 Perawatan Sistem**

Perawatan sistem informasi adalah suatu upaya untuk memperbaiki, menjaga, menanggulangi, mengembangkan sistem yang ada. Perawatan ini diperlukan untuk meningkatkan efisiensi dan efektivitas kinerja sistem yang ada agar dalam penggunaannya dapat optimal. Beberapa alasan mengapa kita perlu memelihara sistem yang ada yaitu: agar dapat meningkatkan sistem / kinerja sistem, dan menyesuaikan dengan perkembangan, agar sistem yang ada tidak tertinggal.

Aplikasi yang professional dalam SDLC dan teknik maupun perangkat modeling yang mendukungnya adalah hal-hal keseluruhan yang terbaik yang dapat seseorang lakukan untuk meningkatkan maintainabilitas sistem.

Jenis-jenis perawatan sistem meliputi :

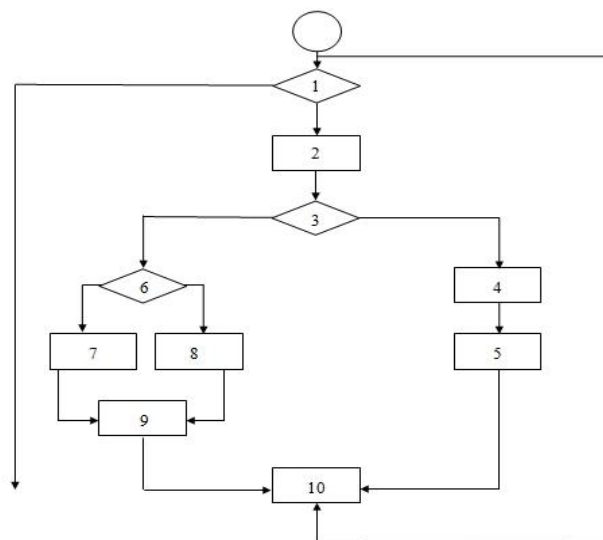
1. Perawatan korektif: adalah pemeliharaan yang mengkoreksi kesalahan-kesalahan yang ditemukan pada sistem, pada saat sistem di jalankan/berjalan.
2. Pemeliharaan adaptif: yaitu pemeliharaan yang bertujuan untuk menyesuaikan perubahan yang terjadi.
3. Pemeliharaan perfektif: pemeliharaan ini bertujuan untuk meningkatkan cara kerja suatu sistem.
4. Pemeliharaan preventif: pemeliharaan ini bertujuan untuk menangani masalah-masalah yang ada.

#### **2.2.11 Teknik Pengujian Sistem**

Pengujian sistem adalah elemen kritis dari jaminan kualitas perangkat lunak dan mempresentasikan kajian pokok dari spesifikasi, desain, dan pengkodean. Tujuan dari pengujian ini adalah diharapkan dengan minimal tenaga dan waktu untuk menemukan berbagai potensi kesalahan dan cacat. Harus didasarkan pada kebutuhan berbagai tahap pengembangan, desain dan dokumen lain atau program yang dirancang untuk menguji struktur internal, dan menggunakan contoh-contoh ini untuk menjalankan program untuk mendeteksi kesalahan. Pengujian sistem informasi harus mencakup pengujian perangkat lunak, pengujian perangkat keras dan pengujian jaringan. Pengujian hardware, jaringan pengujian berdasarkan indikator kinerja spesifik yang akan digunakan di sini pengujian lebih jauh adalah pengujian perangkat lunak.

### 2.2.12 Pengujian Sistem White Box

Pengujian *White box*, adalah metode pengujian yang menggunakan struktur kontrol desain produser untuk test case. Yang memberikan jaminan bahwa semua jalur independen pada suatu modul telah digunakan paling tidak satu kali, menggunakan semua keputusan logis pada sisi true dan false, mengeksekusi semua loop pada batasan mereka dan pada batas operasional mereka, dan menggunakan struktur dan internal untuk menjamin validitasnya.[14]



**Gambar 2.4 Contoh Bagan Alir White Box**

(Sumber: Roger S. Pressman, 2002:536)

Bagan alir digunakan untuk menggambarkan struktur control program dan untuk menggambarkan grafik alir, harus memperhatikan representasi desain procedural pada bagan alir. Pada gambar dibawah ini, grafik alir memetakan bagan alir tersebut ke dalam grafik alir yang sesuai (dengan mengasumsikan bahwa tidak ada kondisi senyawa yang diisikan dalam diamond keputusan bagan alir tersebut). Masing-masing lingkaran, yang disebut simpul grafik alir,





ditemukan. Sebagai contoh, serangkaian jalur independen untuk grafik alir yang ditunjukkan pada gambar di atas adalah:

Jalur 1 : 1 - 11

Jalur 2 : 1 - 2 - 3 - 4 - 5 - 10 - 1 - 11

Jalur 3 : 1 - 2 - 3 - 6 - 8 - 9 - 10 - 1 - 11

Jalur 4 : 1 - 2 - 3 - 6 - 7 - 9 - 10 - 1 - 11

Jalur 1,2,3, dan 4 yang ditentukan diatas terdiri dari sebuah basis set untuk grafik alir pada gambar 2.11. Bagaimana kita tahu banyaknya jalur yang dicari? Komputasi kompleksitas siklomatis memberikan jawaban. Fondasi kompleksitas siklomatis adalah teori grafik, dan memberi kita metrics perangkat lunak yang sangat berguna. Kompleksitas dihitung dalam salah satu dari tiga cara berikut:

1. Jumlah region grafik alir sesuai dengan kompleksitas siklomatis.
2. Kompleksitas siklomatis,  $V(G)$ , untuk grafik  $G$  ditentukan sebagai  $V(G) = E - N + 2$  dimana  $E$  adalah jumlah edge grafik alir dan  $N$  adalah jumlah simpul grafik alir.
3. Kompleksitas siklomatis,  $V(G)$ , untuk grafik alir  $G$  juga ditentukan sebagai  $V(G) = P + 1$ , dimana  $P$  adalah jumlah simpul predikat yang diisikan dalam grafik alir.

Pada gambar tersebut, grafik alir, kompleksitas siklomatis dapat dihitung dengan menggunakan masing-masing dari algoritma yang ditulis diatas:

1. Grafik alir mempunyai 4 region
2.  $V(G) = 11 \text{ edge} - 9 \text{ simpul} + 2 = 4$

3.  $V(G) = 3$  simpul yang diperkirakan  $+ 1 = 4$  Dengan demikian, kompleksitas siklomatis dari grafik alir pada gambar 2.9.
4. Yang lebih penting, nilai untuk  $V(G)$  member kita batas atas untuk jumlah jalur independen yang membentuk basis set, dan implikasinya, batas atas.

### 2.2.13 Pengujian Sistem Black Box

*Black Box Approach* adalah sistem dimana *input* dan *output*nya dapat didefinisikan tetapi prosesnya tidak diketahui atau tidak terdefinisi. Metode ini hanya dapat dimengerti oleh pihak dalam (yang menangani sedangkan pihak luar hanya mengetahui masukan dan hasilnya). Sistem ini terdapat pada subsistem tingkat terendah.

Metode uji coba *Black Box* memfokuskan pada keperluan fungsional dari *software*. Karena uji coba *Black Box* memungkinkan pengembangan *software* untuk membuat himpunan kondisi *input* yang akan melatih seluruh syarat-syarat fungsional suatu program. Uji coba *Black Box*, tetapi merupakan pendekatan yang melengkapi untuk menemukan kesalahan lainnya, selain menggunakan metode *White Box*.

Uji coba *Black Box* berusaha untuk menemukan kesalahan dalam beberapa kategori, diantaranya:

1. Fungsi-fungsi yang salah atau hilang
2. Kesalahan *interface*.
3. Kesalahan dalam struktur data atau akses *database eksternal*.
4. Kesalahan performa.
5. Kesalahan inisialisasi dan terminasi.

Tidak seperti metode *White Box* yang dilaksanakan diawal proses, uji coba *Black Box* diaplikasikan di beberapa tahapan berikutnya. Karena uji coba *Black Box* dengan sengaja mengabaikan struktur kontrol, sehingga perhatiannya difokuskan pada informasi *domain*. Uji coba didesain untuk dapat menjawab pertanyaan-pertanyaan berikut ini:

- (a) Bagaimana validasi fungsional diuji?
- (b) Jenis *Input* apa yang akan menghasilkan kasus uji yang baik?
- (c) Apakah sistem secara khusus sensitif terhadap nilai input tertentu?
- (d) Bagaimana batasan-batasan kelas data diisolasi?
- (e) Berapakah rasio data dan jumlah data yang dapat ditoleransi oleh sistem?
- (f) Apa akibat yang timbul dari kombinasi spesifik data pada operasi sistem?
- (g) Dengan menghasilkan uji coba *Black Box* diharapkan dapat menghasilkan kasus uji coba yang memenuhi kriteria berikut ini:

Kasus uji coba yang berkurang, jika jumlahnya lebih dari 1, maka jumlah dari uji kasus tambahan harus didesain untuk mencapai uji coba yang cukup beralasan.

Kasus uji coba yang memberitahukan sesuatu tentang keberadaan atau tidaknya suatu jenis kesalahan, dari pada kesalahan yang terhubung hanya dengan satu uji coba yang spesifik.

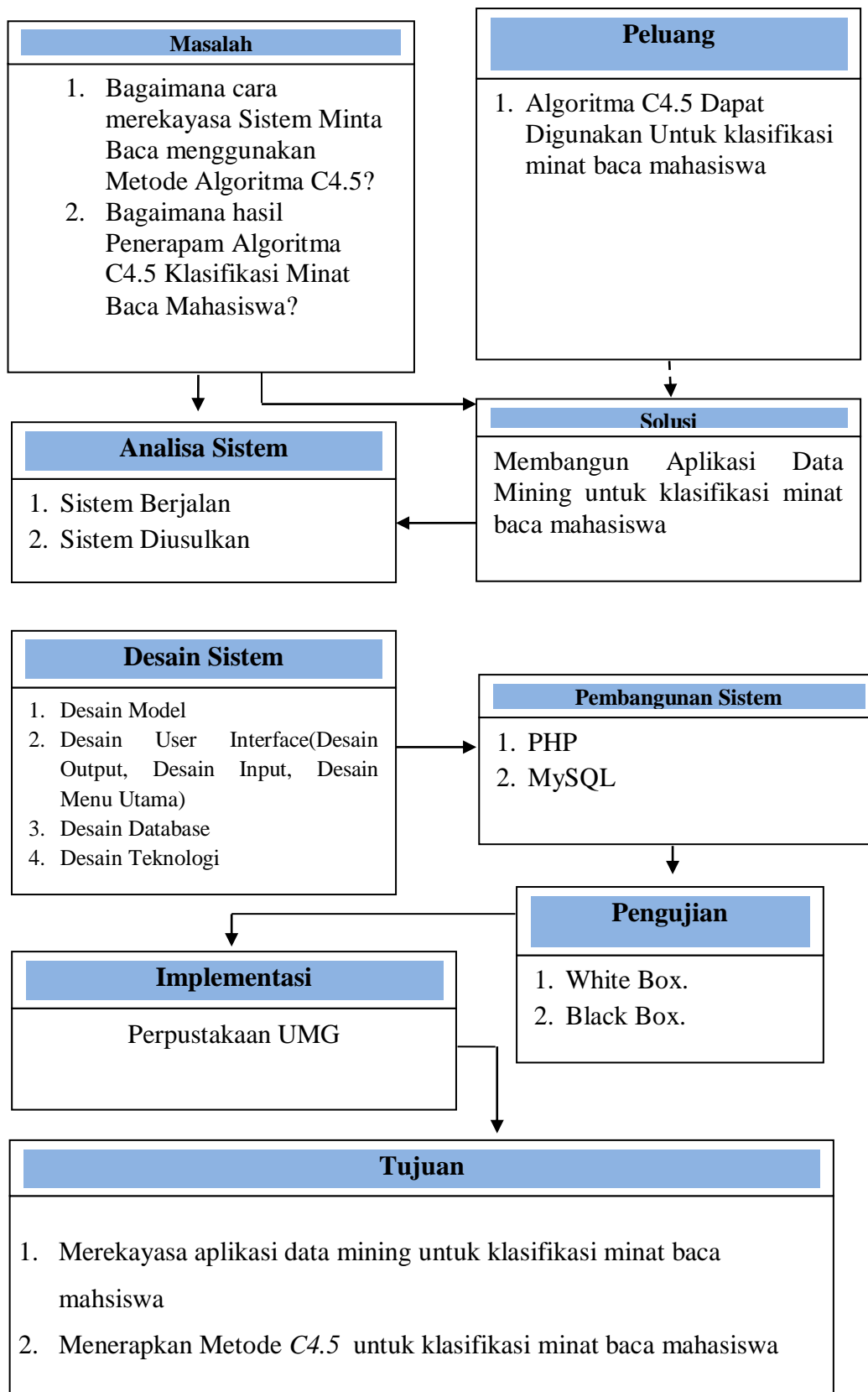
#### 2.2.14 Perangkat Lunak Pendukung

Adapun perangkat lunak pendukung yang digunakan oleh penulis dalam membangun sistem ini ada beberapa diantaranya adalah:

**Tabel 2.5**Perangkat Lunak Pendukung

No.	Perangkat Lunak Pendukung	Kegunaan
1.	PHP	Bahasa Pemrograman yang digunakan untuk membuat program.
2.	Database MySQL	Sebuah perangkat lunak yang digunakan dalam pengoperasian basis data.

### 2.3 Kerang Pikir



**Gambar 2.6** Kerangka Pikir

## **BAB III**

### **METODE PENELITIAN**

#### **3.1 Objek Penelitian**

Penelitian ini menggunakan metode penelitian eksperimen. Objek penelitian ini adalah minat baca, penelitian ini dimulai dari bulan maret 2021 yang berlokasi pada Perpustakaan Universitas Muhammadiyah Gorontalo.

#### **3.2 Metode Penelitian**

Metode yang digunakan dalam penelitian ini adalah metode deskriptif. Metode deskriptif adalah metode yang digunakan untuk menggambarkan atau menganalisis suatu hasil penelitian tetapi tidak digunakan untuk membuat kesimpulan yang lebih luas.[16] Sumber data yang digunakan dalam penelitian ini adalah sumber data sekunder, dimana data yang diperoleh penulis merupakan data yang diperoleh secara tidak langsung, artinya data-data tersebut berupa data primer yang telah diolah lebih lanjut dan data yang disajikan oleh pihak lain.

##### **3.2.1 Pengumpulan Data**

Data prediksi minat baca dikumpulkan menggunakan teknik observasi, kuesioner. Adapun variabel/atribut dengan tipe datanya masing-masing ditunjukkan pada tabel berikut ini.

**Tabel 3.1 Atribut Data Prediksi Minat Baca**

No.	Nama	Type	KET
1.	Jenis Kelamin	Kategorikal	Variabel Input
2.	Usia	Kategorikal	Variabel Input
3.	Jenis Bacaan	Kategorikal	Variabel Input
4.	Tujuan Membaca	Kategorikal	Variabel Input
5.	Lingkungan	Kategorikal	Variabel Input
6.	Minat Baca	Kategorikal	Variabel Output

Data klasifikasi minat baca ditunjukkan pada tabel berikut ini.

**Tabel 3.2 Data Klasifikasi Minat Baca**

Jenis Kelamin	Usia	Jenis Bacaan	Tujuan Membaca	Lingkungan	Minat Baca
L	Remaja	Fiksi	Menambah Wawasan	Dosen	Sedang
P	Dewasa	Non Fiksi	Motivasi	Teman	Tinggi
L	Tua	Fiksi	Menambah Wawasan	Asrama	Tinggi
P	Remaja	Non Fiksi	Motivasi	Dosen	Sedang
L	Dewasa	Fiksi	Menambah Wawasan	Teman	Sedang
P	Tua	Non Fiksi	Motivasi	Dosen	Tinggi

### 3.2.2 Pra Pengolahan

Sebelum data diolah, terlebih dahulu dilakukan normalisasi dan *transformasi data*. Hal ini dilakukan agar data tersebut dapat diolah dengan baik oleh algoritma C.45

### 3.2.3 Validasi Prediksi

Validasi dilakukan membagi data training dan data testing. Teknik yang digunakan pada tahap ini adalah *hold out*.

Prediksi Algoritma C4.5 dilatih menggunakan data training yang telah dikumpulkan. Hal ini dilakukan sebelum model diuji, agar model tersebut dapat memprediksi minat baca mahasiswa dengan akurat.

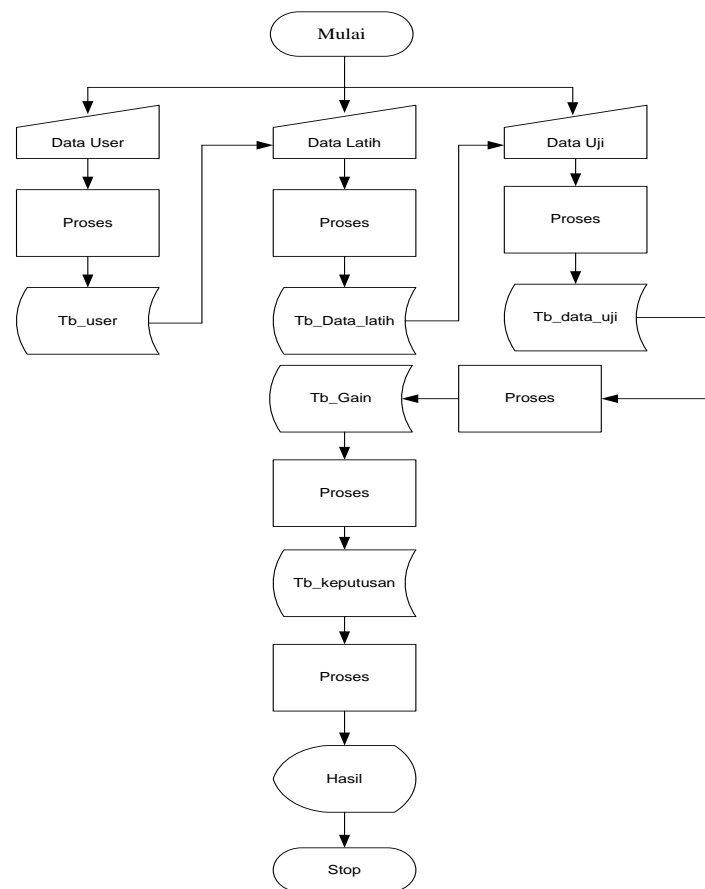
### 3.2.4 Evaluasi Model

Model yang telah dihasilkan kemudian dievaluasi linerjanya dengan menggunakan *confusion Matrix* untuk mengetahui akurasi.

## 3.3 Pengembangan Sistem

Sistem yang diusulkan dapat digambarkan menggunakan *flowchart* sistem yang ditunjukkan pada Gambar 3.1 berikut ini.





**Gambar 3.1 Sistem yang Diusulkan**

### 3.3.1 Tahap Desain

Pada tahap ini dilakukan desain sistem yakni desain output, desain input, desain database, desain teknologi dan desain model :

#### a. Desain Output

Pada tahap ini dilakukan desain output secara umum dan terinci yakni desain output laporan hasil prediksi dan hasil akurasi.

#### b. Desain Input

Pada tahap ini dilakukan desain input secara umum dan terinci, yakni desain input ,dan dataset.

#### c. Desain Database

Pada tahap ini dilakukan desain database yang dimaksudkan untuk mendefinisikan isi atau struktur dari tiap-tiap file yang telah diidentifikasi didesain secara umum.

d. Desain Teknologi

Pada tahap ini kita menentukan teknologi yang akan dipergunakan dalam menerima input, menjalankan model, menyimpan dan mengakses data, menghasilkan dan mengirimkan keluaran dan membantu pengendalian dari sistem secara keseluruhan. Teknologi yang dimaksud meliputi perangkat keras, perangkat lunak yang akan digunakan serta sumber daya manusia yang akan menggunakan sistem ini nantinya.

e. Desain Model

Pada tahap ini dilakukan desain model secara umum berupa desain sistem secara fisik dan logika. Desain fisik dapat digambarkan dengan bagan alir sistem dan bagan alir dokumen. Desain secara logika digambarkan dengan diagram arus data (DAD). Pada tahap desain model terinci, model akan mendefinisikan secara rinci urutan-urutan langkah dari masing-masing proses yang digambarkan di DAD.

### 3.3.2 Tahap Produksi atau Pembuatan

Pada tahap ini dilakukan pembuatan sistem dengan menggunakan Bahasa Pemrograman *PHP* dengan memanfaatkan Database *MySQL*. Pada tahap ini kita melakukan tahap produksi sistem hasil analisa dan desain sistem sebelumnya. Termasuk didalamnya menginstal paket tambahan untuk menjalankan program, menulis listing program dan membangunnya dalam bentuk sebuah formulir,

antarmuka dan integrasi sistem-sistem program yang terdiri dari input, proses dan output yang tersusun dalam sebuah sistem menu sehingga dapat dijalankan oleh pengguna sistem.

### **3.3.3 Tahap Pengujian**

Setelah dilakukan tahap analisa, desain dan produksi sistem, maka kita melakukan tahap pengujian, dimana seluruh perangkat lunak, program tambahan dan semua program yang terlibat dalam pembangunan sistem diuji untuk memastikan sistem dapat berjalan dengan semestinya. Testing difokuskan pada logika internal, fungsi eksternal dan mencari segala kemungkinan kesalahan dari sistem yang dibuat. Pada tahap ini dilakukan review dan evaluasi terhadap sistem yang dikembangkan, apakah sudah sesuai dengan rancangan atau belum. Jika terjadi hal-hal yang tidak sesuai dengan yang diharapkan, kemudian dilakukan revisi atau perbaikan supaya produk tersebut dapat dioperasikan dengan baik dan siap untuk diimplementasikan. Pengujian yang dilakukan dengan menggunakan teknik pengujian perangkat lunak yaitu :

- a. Pengujian *White Box* terhadap sistem yang akan digunakan.
- b. Pengujian *Black Box* melalui program *PHP* dan Database *MySQL*.

Setelah dilakukan uji coba sistem secara internal, kemudian dilakukan pengujian antarmuka sistem, apakah sebuah sistem setelah diberikan ke pengguna dapat dioperasikan atau tidak.

### 3.3.4 Pengujian Sistem

#### a. White Box Testing

Software yang telah direkayasa kemudian diuji dengan metode *White Box Testing* pada kode program proses penerapan metode/modelnya. Kode program tersebut dibuatkan *flowchart* programnya, kemudian dipetakan kedalam bentuk *flowgraph* (bagan alir kontrol) yang tersusun dari beberapa *node* dan *edge*. Berdasarkan *flowgraph*, ditentukan jumlah *Region* dan *Cyclomatic Complexity* (CC). apabila  $independent\ path = V(G) = (CC) = Region$ , di mana setiap *path* hanya dieksekusi sekali dan sudah benar, maka sistem dinyatakan efisien dari segi kelayakan logika pemrograman.

#### b. Black Box Testing

Selanjutnya *software* diuji pula dengan metode *Black Box Testing* yang focus pada keperluan fungsional dari *software* dan berusaha untuk menemukan kesalahan dalam beberapa kategori, diantaranya:

1. Fungsi-fungsi yang salah atau hilang.
2. Kesalahan interface.
3. Kesalahan dalam struktur data atau akses basis data external.
4. Kesalahan performa.
5. Kesalahan inisialisasi dan terminasi.

Jika sudah tidak ada kesalahan-kesalahan tersebut, maka sistem dinyatakan efisien dari segi kesalahan komponen-komponen sistem.

## BAB IV

### HASIL PENELITIAN

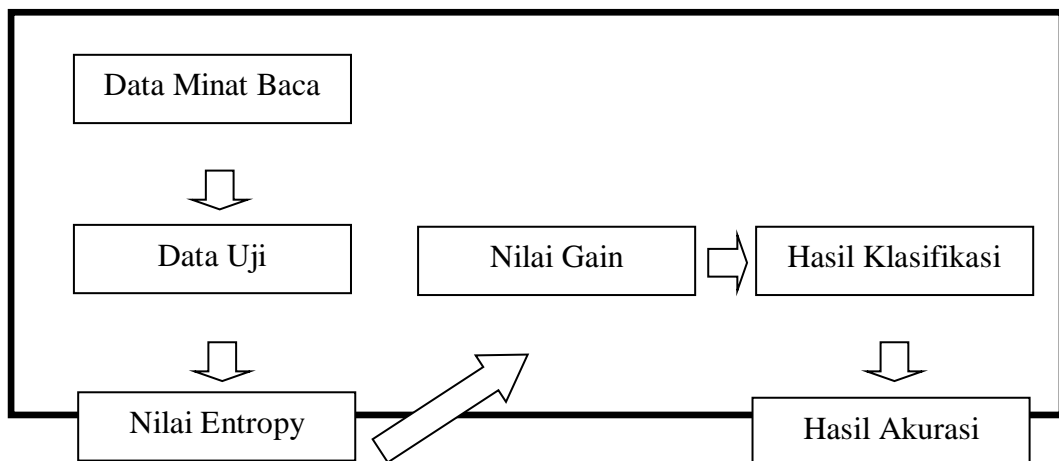
#### 4.1 Hasil Pengumpulan Data

**Tabel 4.1** Data Penelitian

No	Umur	Jenis Kelamin	Jenis Bacaan	Tujuan Membaca	Lingkungan Baca	Minat Membaca
1	41	Wanita	Non Fiksi	Menambah Wawasan	Alam Terbuka	Sedang
2	47	Wanita	Fiksi	Menambah Wawasan	Didalam ruangan	Sedang
3	23	Pria	Fiksi	Menambah Wawasan	Alam Terbuka	Sedang
4	27	Pria	Non Fiksi	Menambah Wawasan	Didalam ruangan	Sedang
5	18	Wanita	Non Fiksi	Menambah Wawasan	Didalam ruangan	Tinggi
6	19	Wanita	Non Fiksi	Menambah Wawasan	Didalam ruangan	Sedang
7	19	Wanita	Non Fiksi	Menambah Wawasan	Didalam ruangan	Sedang
8	18	Wanita	Non Fiksi	Menambah Wawasan	Didalam ruangan	Sedang
9	20	Wanita	Fiksi	Motivasi	Didalam ruangan	Sedang
10	20	Wanita	Fiksi	Menambah Wawasan	Didalam ruangan	Sedang
11	20	Wanita	Non Fiksi	Menambah Wawasan	Didalam ruangan	Sedang
12	20	Wanita	Non Fiksi	Menambah Wawasan	Didalam ruangan	Sedang
13	19	Pria	Fiksi	Motivasi	Alam Terbuka	Tinggi
14	19	Wanita	Fiksi	Menambah Wawasan	Alam Terbuka	Tinggi
15	20	Wanita	Non Fiksi	Menambah Wawasan	Didalam ruangan	Tinggi
16	20	Wanita	Fiksi	Menambah Wawasan	Alam Terbuka	Sedang

17	19	Wanita	Non Fiksi	Motivasi	Alam Terbuka	Sedang
18	19	Pria	Fiksi	Motivasi	Alam Terbuka	Sedang
19	21	Wanita	Fiksi	Menambah Wawasan	Didalam ruangan	Sedang
20	20	Pria	Fiksi	Motivasi	Alam Terbuka	Sedang
.....	.....	.....	.....	.....	.....	.....
.....	.....	.....	.....	.....	.....	.....
200	27	Wanita	Fiksi	Motivasi	Didalam ruangan	Tinggi
201	35	Wanita	Non Fiksi	Menambah Wawasan	Didalam ruangan	Tinggi

## 4.2 Arsitektur Algoritma C4.5



**Gambar 4.1** Arsitektur Algoritma C4.5

### 4.3 Penerapan Metode C4.5

**Tabel 4.2** Data Penerapan

No	Usia	Jenis Kelamin	Jenis Bacaan	Tujuan Membaca	Lingkungan Baca	Minat Membaca
1	38	L	Fiksi	Menambah Wawasan	Teman	?

- A. Menghitung jumlah kasus, jumlah kasus untuk keputusan Sedang, jumlah kasus untuk keputusan Rendah, dan Entropy dari semua kasus dan kasus yang dibagi berdasarkan atribut Umur, Jenis Kelamin, Tujuan Membaca, dan Lingkungan Baca.
- B. Pada tahap berikut akan dilakukan perhitungan Entropy Dan Gain untuk setiap atribut.

**Tabel 4.3** Perhitungan Gain

Atribut		Jumlah	Sedang	Tinggi	Entropy	Gain
Total		201	140	61	0,8855064 18	
						0,000150247
Jenis Kelamin						
	L	48	34	14	0,8708644 69	
	P	153	106	47	0,8899025 86	
Usia						0,031465967
	Remaja	158	112	46	0,8701883 35	
	Dewasa	29	15	14	0,9991421 04	
	Tua	14	13	1	0,3712323 27	
Jenis Bacaan						0,163310728
	Fiksi	114	78	36	0,8997437 59	
	Non Fiksi	87	62	78	0,4895464 95	
Tujuan Membaca						-0,01127448

	Motivasai	55	34	21	0,9593160 32	
	Menambah Wawasan	146	102	40	0,8732231 42	
Lingkungan						0,004068694
	Asrama	82	57	25	0,8871723 03	
	Teman	105	75	30	0,8631205 69	
	Dosen	14	8	6	0,9852281 36	

Baris total kolom Entropy pada table 4.4 dihitung dengan persamaan berikut:

$$\text{Entropy (Total)} = \left( -\frac{141}{201} * \text{Log}_2\left(\frac{141}{201}\right) \right) + \left( -\frac{55}{201} * \text{Log}_2\left(\frac{55}{201}\right) \right)$$

$$= 0,879459$$

$$\text{Entropy (Jenis Kelamin, L)} = \left( -\frac{35}{48} * \text{Log}_2\left(\frac{35}{48}\right) \right) + \left( -\frac{13}{48} * \text{Log}_2\left(\frac{13}{48}\right) \right)$$

$$= 0,842658$$

$$\text{Entropy (Jenis Kelamin, P)} = \left( -\frac{105}{153} * \text{Log}_2\left(\frac{105}{153}\right) \right) + \left( -\frac{47}{153} * \text{Log}_2\left(\frac{47}{153}\right) \right)$$

$$= 0,870188335$$

$$\text{Entropy (Usia, Remaja)} = \left( -\frac{112}{158} * \text{Log}_2\left(\frac{112}{158}\right) \right) + \left( -\frac{46}{158} * \text{Log}_2\left(\frac{46}{158}\right) \right)$$

$$= 0,879459$$

$$\text{Entropy (Usia, Dewasa)} = \left( -\frac{15}{29} * \text{Log}_2\left(\frac{15}{29}\right) \right) + \left( -\frac{14}{29} * \text{Log}_2\left(\frac{14}{29}\right) \right)$$

$$= 0,999142104$$

$$\text{Entropy (Usia, Tua)} = \left( -\frac{13}{14} * \text{Log}_2\left(\frac{13}{14}\right) \right) + \left( -\frac{1}{14} * \text{Log}_2\left(\frac{1}{14}\right) \right)$$

$$= 0,031465967$$



$$\text{Entrophy (Jenis Bacaan, Fiksi)} = \left( -\frac{78}{114} * \text{Log}_2\left(\frac{78}{114}\right) \right) + \left( -\frac{36}{114} * \text{Log}_2\left(\frac{36}{114}\right) \right)$$

$$= 0,899743759$$

$$\text{Entrophy (Jenis Bacaan, Non Fiksi)} = \left( -\frac{62}{87} * \text{Log}_2\left(\frac{62}{87}\right) \right) + \left( -\frac{78}{87} * \text{Log}_2\left(\frac{78}{87}\right) \right)$$

$$= 0,489546495$$

$$\text{Entrophy (Tujuan Membaca, Motivasi)} = \left( -\frac{34}{55} * \text{Log}_2\left(\frac{34}{55}\right) \right) + \left( -\frac{21}{55} * \text{Log}_2\left(\frac{21}{55}\right) \right)$$

$$= 0,959316032$$

$$\text{Entrophy (Tujuan Membaca, Menambah Wawasan)} = \left( -\frac{102}{146} * \text{Log}_2\left(\frac{102}{146}\right) \right) +$$

$$\left( -\frac{40}{146} * \text{Log}_2\left(\frac{40}{146}\right) \right) = 0,873223142$$

$$\text{Entrophy (Lingkungan, Asrama)} = \left( -\frac{57}{87} * \text{Log}_2\left(\frac{57}{87}\right) \right) +$$

$$\left( -\frac{25}{87} * \text{Log}_2\left(\frac{25}{87}\right) \right) = 0,887172303$$

$$\text{Entrophy (Lingkungan, Teman)} = \left( -\frac{75}{105} * \text{Log}_2\left(\frac{75}{105}\right) \right) +$$

$$\left( -\frac{30}{105} * \text{Log}_2\left(\frac{30}{105}\right) \right) = 0,863120569$$

$$\text{Entrophy (Lingkungan, Dosen)} = \left( -\frac{8}{14} * \text{Log}_2\left(\frac{8}{14}\right) \right) +$$

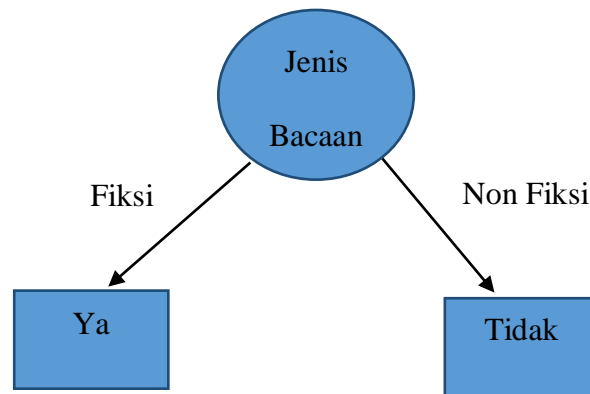
$$\left( -\frac{6}{14} * \text{Log}_2\left(\frac{6}{14}\right) \right) = 0,985228136$$

$$\begin{aligned} \text{Perhitungan Gain Jenis Kelamin} &= (0,885506418) - ((48/201)* \\ &\quad (0,870864469) + (153/201)* \\ &\quad (0,889902586)) \\ &= 0,000150247 \end{aligned}$$

$$\begin{aligned} \text{Perhitungan Gain Usia} &= (0,870188335) - ((158/201)*(0,999142104) \\ &\quad + (112/201)* \\ &\quad (0,889902586)) \\ &= 0,000150247 \end{aligned}$$

Setelah didapat hasil dari tabel 4.1 di atas maka cari nilai Gain yang tertinggi.

Nilai gain tertinggi maka itulah yang menjadi akarnya.

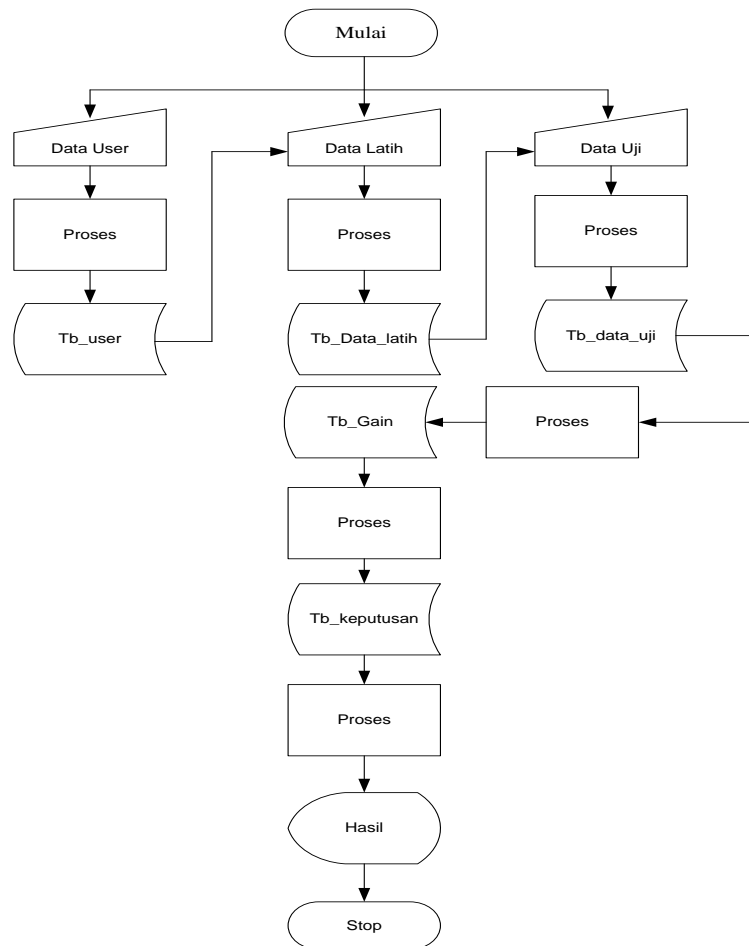


**Gambar 4.2** Root Dari Hasil Gain

Ket:

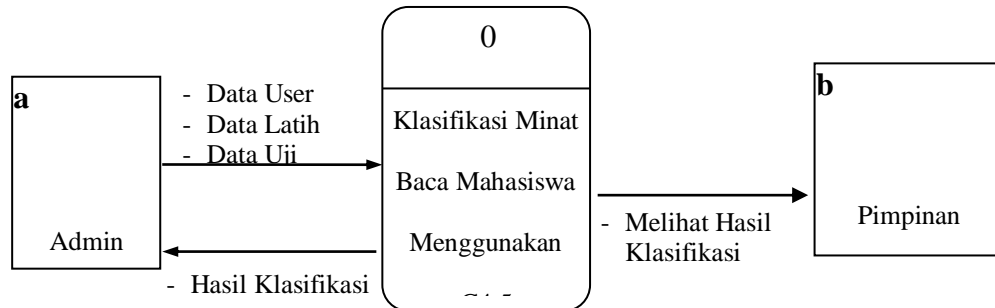
- Jenis Bacaan merupakan gain terbesar, maka Jenis Bacaan dijadikan sebuah akar
- Pada tabel 4.3 dapat kita lihat bahwa pada Jenis Bacaan yang menyatakan Fiksi semuanya terletak pada kolom tidak sedangkan pada kolom ya bernilai lebih rendah. Maka dipastikan keluaran dari Jenis Bacaan-Fiksi yaitu tidak

#### 4.4 Sistem Yang Diusulkan



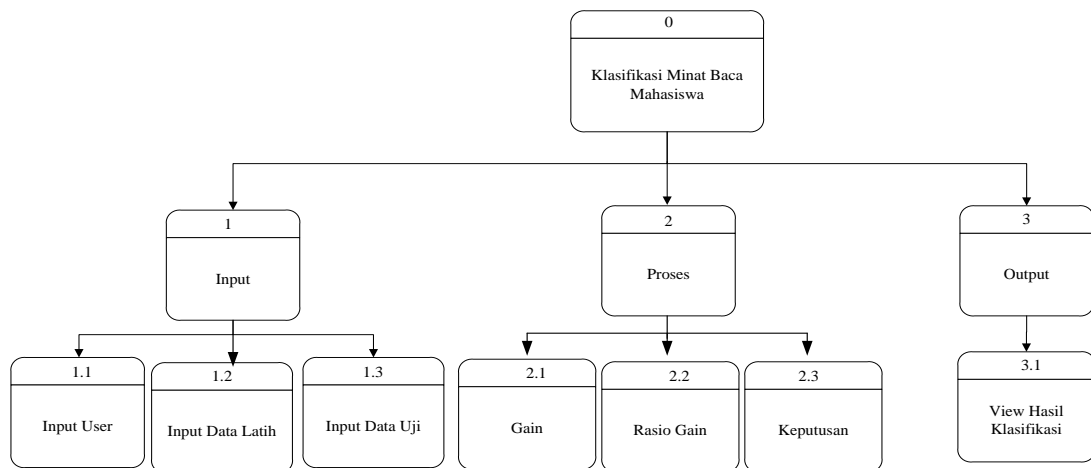
**Gambar 4.3** Sistem yang diusulkan

#### 4.5 Diagram Konteks



**Gambar 4.4**Diagram Konteks

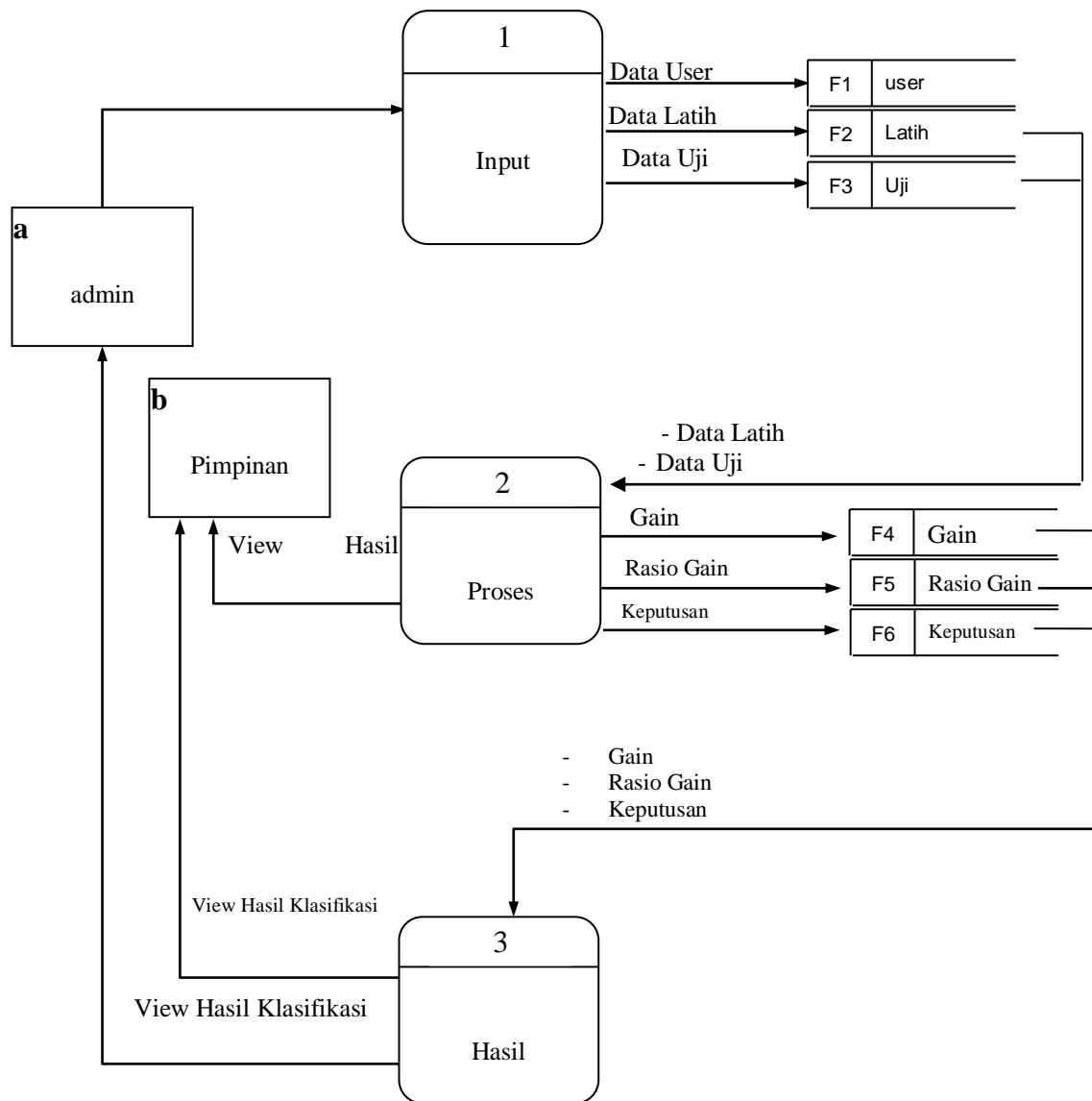
#### 4.6 Diagram Berjenjang



**Gambar 4.5**Diagram Berjenjang

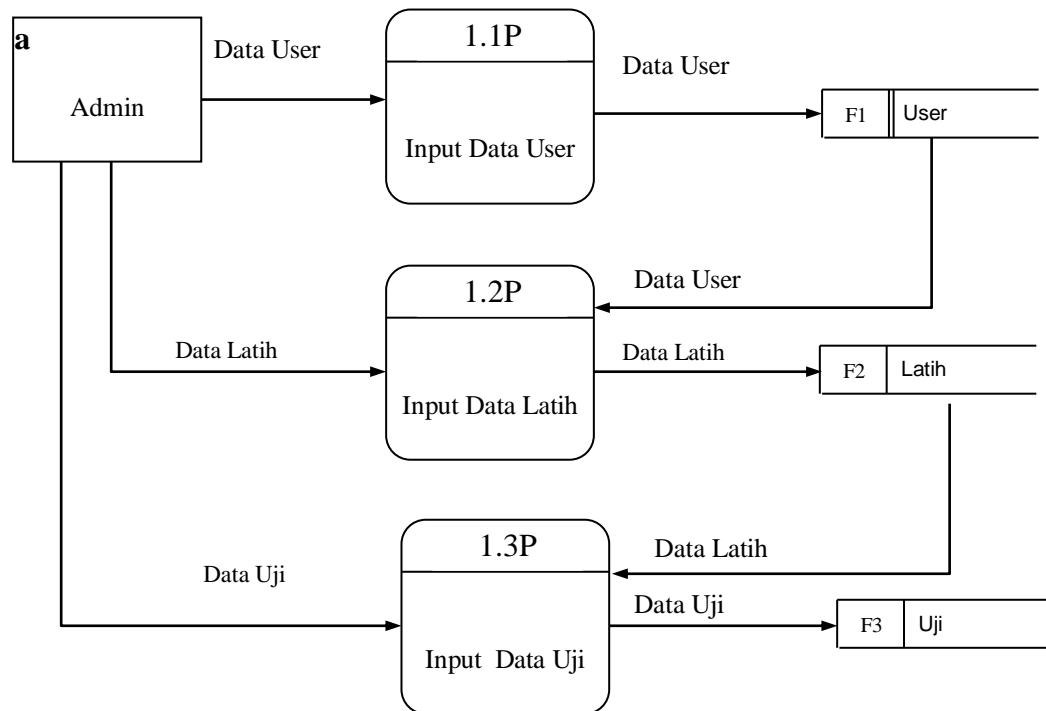
## 4.7 Diagram Arus Data

### 4.7.1 DAD Level 0



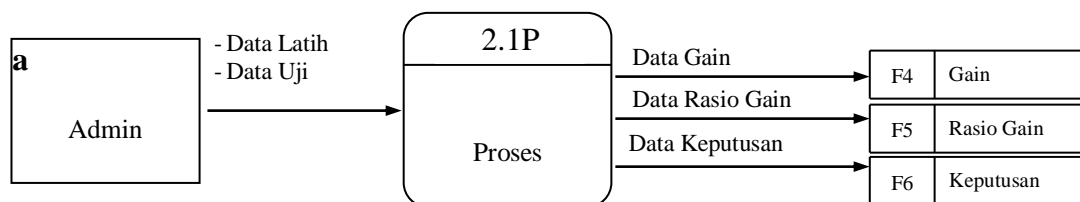
**Gambar 4.6** DAD Level 0

#### 4.7.2 DAD Level 1 Proses 1



**Gambar 4.7**DAD Level 1 Proses 1

#### 4.7.3 DAD Level 1 Proses 2



**Gambar 4.8**DAD Level 1 Proses 2

#### 4.7.4 DAD Level 1 Proses 3



**Gambar 4.9**DAD Level 1 Proses 3

#### 4.8. Kamus Data

##### 4.8.1 Kamus Data User

**Tabel 4.4.** Kamus Data User

Kamus Data : Detail Data User				
Nama Arus Data : Data Detail User			Bentuk Data :	
Penjelasan : Berisi data-data detail User untuk penginputan data			Dokumen, Laporan	
Periode : Non periodik			Arus Data :a-1,1-	
Struktur Data sebagai Berikut :			F1,F1-2,a-1.1P,1.1P-	
			F1,F1-1.2P	
No	Nama Item Data	Type	Width	Keterangan
1	Id_User	Int	4	Id User
2	Nama	Varchar	100	Nama
3	Username	Varchar	15	Username
4	Password	Varchar	8	Password
5	Level	Varchar	100	Level

#### 4.8.2 Kamus Data Latih

**Tabel 4.5** Kamus Data Latih

<b>Kamus Data : Detail Data Latih</b>				
Nama Arus Data : Data Latih Penjelasan : Berisi data-data untuk penginputan data Latih Periode : Non periodik Struktur Data sebagai Berikut :			Bentuk Data : Dokumen, Laporan Arus Data : a-1,1-F2,F2-2,a-1.2P,1.2P-F2,F2-1.3P	
No	Nama Item Data	Type	Width	Keterangan
1	Id	Int	4	Id
2	Nama	Varchar	100	Nama
3	Umur	Varchar	100	Umur
4	Jeniskelamin	Varchar	100	Jenis Kelamin
5	Jenisbacaan	Varchar	100	Jenis Bacaan
6	Tujuanmembaca	Varchar	100	Tujuan Membaca
7	Lingkunganbaca	Varchar	100	Lingkungan Baca
8	Minatmembaca	Varchar	100	Minat Membaca



### 4.8.3 Kamus Data Uji

**Tabel 4.6** Kamus Data Uji

<b>Kamus Data : Detail Data Uji</b>				
Nama Arus Data : Data Uji Penjelasan : Berisi data-data untuk penginputan data Uji Periode : Non periodik Struktur Data sebagai Berikut :			Bentuk Data : Dokumen, Laporan Arus Data : a-1,1-F3,F3-2,a-1.3P,1.3P-F3	
No	Nama Item Data	Type	Width	Keterangan
1	Id	Int	4	Id
2	Nama	Varchar	100	Nama
3	Umur	Varchar	100	Umur
4	Jeniskelamin	Varchar	100	Jenis Kelamin
5	Jenisbacaan	Varchar	100	Jenis Bacaan
6	Tujuanmembaca	Varchar	100	Tujuan Membaca
7	Lingkunganbaca	Varchar	100	Lingkungan Baca

#### 4.8.4 Kamus Data Gain

**Tabel 4.7** Kamus Data Gain

<b>Kamus Data : Detail Data Gain</b>				
Nama Arus Data : Data Gain Penjelasan : Berisi data-data untuk penginputan data Gain Periode : Non periodik Struktur Data sebagai Berikut :			Bentuk Data : Dokumen, Laporan Arus Data : F4-3,F4-3.1P	
No	Nama Item Data	Type	Width	Keterangan
1	Id	Int	4	Id
2	node_id	Varchar	100	Node Id
3	Attribute	Varchar	100	Attribut
4	Gain	Varchar	100	Gain

#### 4.8.5 Kamus Data Keputusan

**Tabel 4.8.** Kamus Data Keputusan

<b>Kamus Data : Detail Keputusan</b>				
Nama Arus Data : Data Keputusan Penjelasan : Berisi data-data untuk penginputan data Keputusan Periode : Non periodik Struktur Data sebagai Berikut :			Bentuk Data : Dokumen, Laporan Arus Data : F6-3,F6-3.1P	
No	Nama Item Data	Type	Width	Keterangan
1	Id	Int	11	Id
2	Parent	Varchar	100	Parent
3	Akar	Varchar	100	Akar
4	Keputusan	Varchar	100	Keputusan

#### 4.8.6 Kamus Data Hasil Klasifikasi

**Tabel 4.9.** Kamus Data Hasil Klasifikasi

<b>Kamus Data : Detail Hasil Klasifikasi</b>				
Nama Arus Data : Data Klasifikasi Penjelasan : Berisi data-data untuk penginputan data Klasifikasi Periode : Non periodik Struktur Data sebagai Berikut :			Bentuk Data : Dokumen, Laporan Arus Data :F6-3,F6-3.1P	
No	Nama Item Data	Type	Width	Keterangan
1	Id	Int	4	Id
2	Nama	Varchar	100	Nama
3	Umur	Varchar	100	Umur
4	Jeniskelamin	Varchar	100	Jenis Kelamin
5	Jenisbacaan	Varchar	100	Jenis Bacaan
6	Tujuanmembaca	Varchar	100	Tujuan Membaca
7	Lingkunganbaca	Varchar	100	Lingkungan Baca
8	Hasilklasifikasi minatmembaca	Varchar	100	Hasil Klasifikasi Minat Membaca
9	Id_rule	Int	11	Id Rule

#### 4.9 Arsitektur Klasifikasi Peminat Baca

Sistem Klasifikasi menggunakan jaringan client server. Sedangkan spesifikasi hardware dan software yang direkomendasikan, yaitu :

- Processor : Intel Celeron – Intel Core i7
- RAM : 1 GB
- VGA : 1024 pixel
- Harddisk : 250GB
- Operating System : Windows 7 – windows 10
- Tools : Notepad++, Xampp, Google Chrome

#### 4.10 Interface Desain Mekanisme User

**Tabel 4.10**Mekanisme User

User	Kategori	Akses Input	Akses Output
Admin	Administrator	All	All
User	User	-	Hasil Klasifikasi

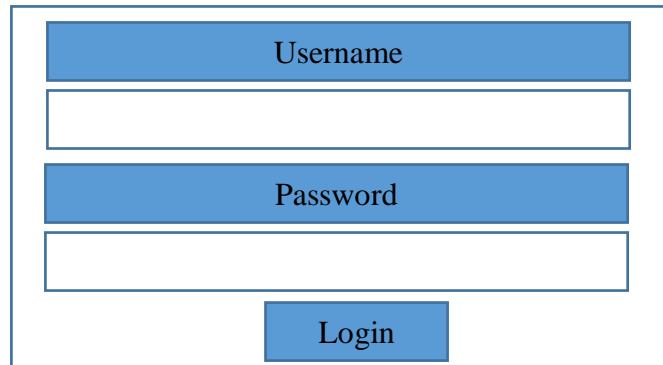
#### 4.11 Intrface Desain



**Gambar 4.10** Interface Desain

## 4.12 Mekanisme Input

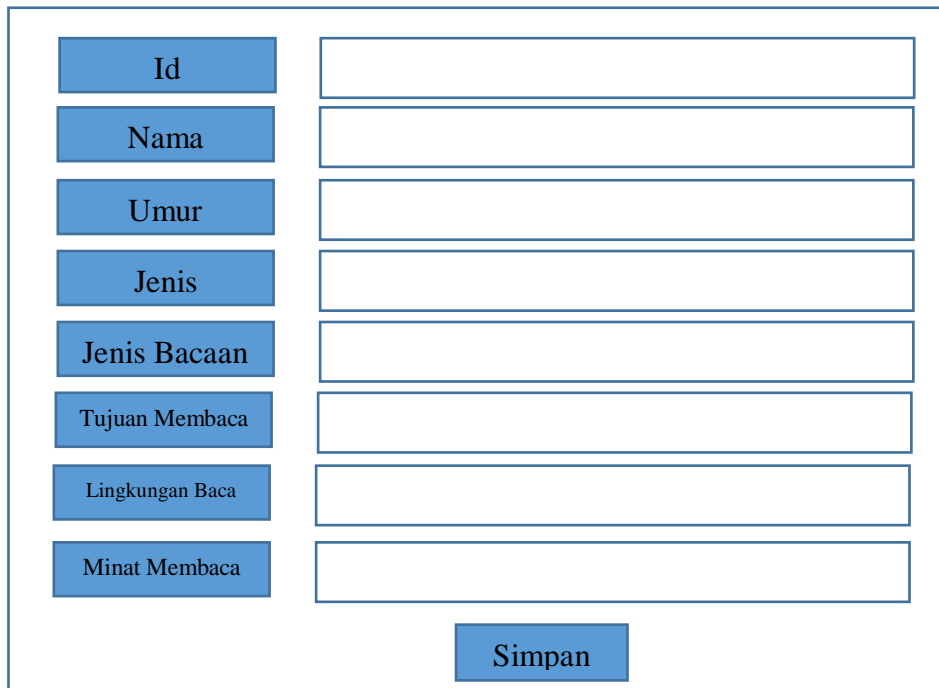
### 1. Input Login Admin



A login form interface with a blue border. It contains two input fields: 'Username' and 'Password', each with a blue label above a white input box. Below the input fields is a blue 'Login' button.

**Gambar 4.11** Interface Login Admin

### 2. Input Data Latih



A data entry form interface with a blue border. It features eight labels in blue boxes on the left, each corresponding to a white input field on the right: 'Id', 'Nama', 'Umur', 'Jenis', 'Jenis Bacaan', 'Tujuan Membaca', 'Lingkungan Baca', and 'Minat Membaca'. At the bottom center is a blue 'Simpan' button.

**Gambar 4.12** Interface Data Latih

### 3. Input Data uji

Id	<input type="text"/>
Nama	<input type="text"/>
Umur	<input type="text"/>
Jenis	<input type="text"/>
Jenis Bacaan	<input type="text"/>
Tujuan Membaca	<input type="text"/>
Lingkungan Baca	<input type="text"/>
<input type="button" value="Proses"/>	

**Gambar 4.13** Interface Data Uji

### 4.13 Interface Output

Header	
Navigasi Menu	Hasil Klasifikasi

**Gambar 4.14** Interface Output

#### 4.14 Data Desain

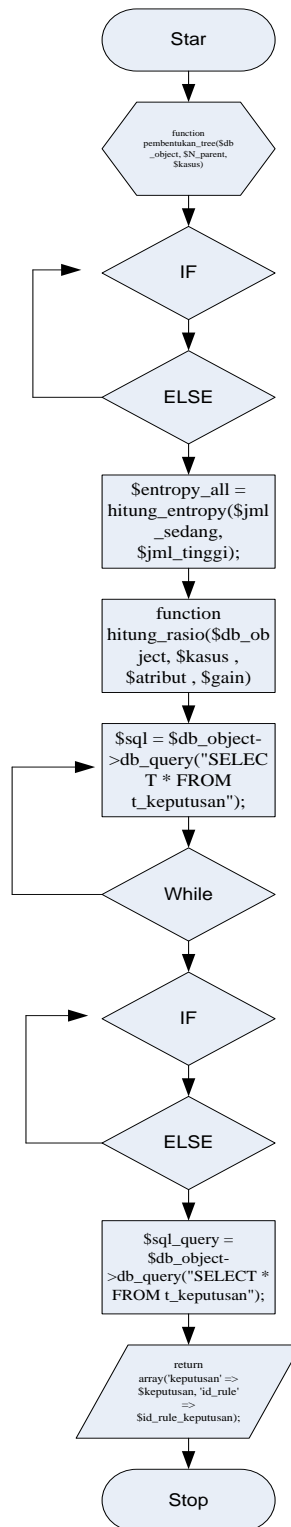
Data yang diperoleh pada sistem Prediksi Alat Kontrasepsi ini menggunakan format

- **Notepad++** sebagai Editor Coding
- Dataset **Mysql server** untuk mengolah dan menyimpan data
- Keduanya dihubungkan dan dimanipulasi dengan teknik disconnected data

#### 4.15 Pengujian WhiteBox

```
function pembentukan_tree($db_object, $N_parent, $kasus)..... 1
if ($N_parent != ""){ $kondisi = $N_parent . " AND " . $kasus; } ..... 2
else { $kondisi = $kasus; } ..... 3
$entropy_all = hitung_entropy($jml_sedang, $jml_tinggi); ..... 4
function hitung_rasio($db_object, $kasus , $atribut , $gain) ..... 5
$sql = $db_object->db_query("SELECT * FROM t_keputusan"); ..... 6
while ($row = $db_object->db_fetch_array($sql)) { ..... 7
if ($row['parent'] != "") { $rule = $row['parent'] . " AND " . $row['akar'];}.....8
else { $rule = $row['akar']; } ..... 9
$sql_query = $db_object->db_query("SELECT * FROM t_keputusan"); ..... 10
return array('keputusan' => $keputusan, 'id_rule' => $id_rule_keputusan);..... 11
```

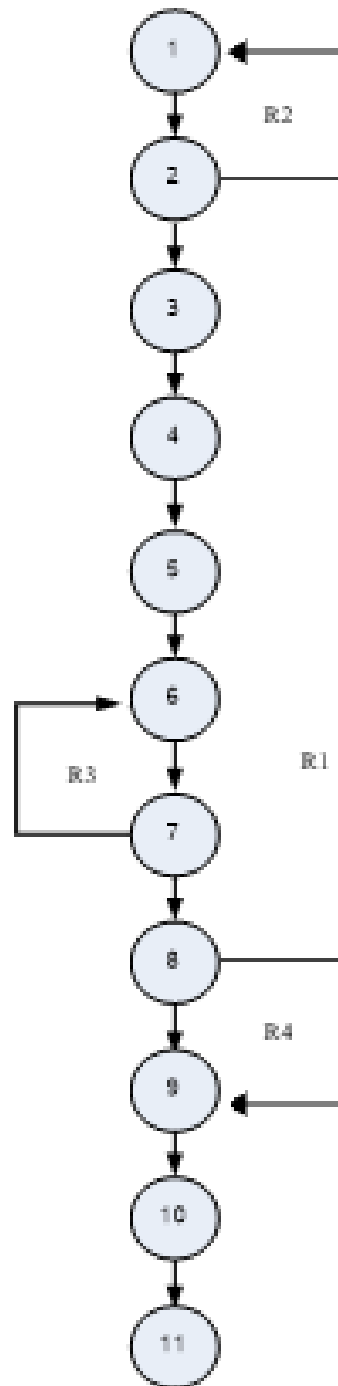
#### 4.16 Flowchart Pengujian WhiteBox



**Gambar 4.15**Flowchart



#### 4.17 Flowgraph



**Gambar 4.16** Flowgraph

#### 4.18 Perhitungan Pengujian White Box

Diketahui :

$$\text{Region (R)} = 4$$

$$\text{Node (N)} = 11$$

$$\text{Edge (E)} = 13$$

$$\text{Predikat Node (P)} = 3$$

$$\text{Rumus : } V(G) = (E-N)+2$$

$$\text{Atau : } V(G) = P + 1$$

$$\text{Penyelesaian : } V(G) = 13 - 11 + 2 = 4$$

$$V(G) = 3 + 1 = 4$$

$$(R1, R2, R3, R4)$$

#### 4.19 Pengujian Basis Path

**Tabel 4.11** Pengujian Path

NO	PATH	KETERANGAN
1	1-2-3-2.....11	OK
2	1-2-3-4-5-6-7-6.....11	OK
3	1-2-3-4-5-6-7-8-9-8.....11	OK
4	1-2-3-4-5-6-7-8-9-10-11	OK

#### 4.20 Pengujian BlackBox

**Tabel 4.12** Hasil Pengujian BlackBox

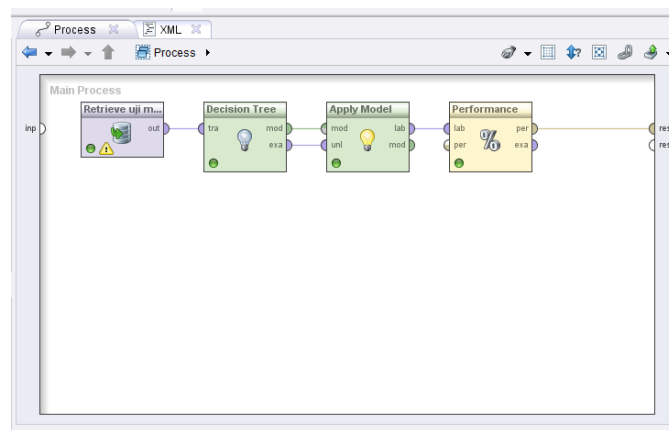
No	Input/Event	Fungsi	Hasil	Hasil
1	Login	Login dengan menginput username dan password Lalu enter	A.Jika username dan password salah maka Kembali Kehalaman Index B.Jika username dan password benar maka akan masuk ke window utam	Sesuai
2	Menu Utama	Menampilkan Halaman Utama	Halaman Utama tampil dan aktif	Sesuai
3	Menu Data Latih	Menampilkan Halaman Data Latih	Halaman dataset tampil dan aktif	Sesuai
4	Input Data Latih	Menampilkan Halaman penginputan data Latih	Halaman penginputan Latih tampil dan aktif	Sesuai
5	Menu Data Uji	Menampilkan Halaman Data Uji	Halaman DataUji tampil dan aktif	Sesuai
6	Input Data Uji	Menampilkan Halaman penginputan Data Uji	Halaman penginputan Data Uji tampil dan aktif	Sesuai
7	Menu Hasil Klasifikasi	Menampilkan Halaman Hasil Klasifikasi	Halaman Hasil Klasifikasi tampil dan aktif	Sesuai
8	Menu User	Menampilkan Halaman user	Halaman user tampil dan aktif	Sesuai
9	Input User	Menampilkan Halaman penginputan User	Halaman penginputan User tampil dan aktif	Sesuai
10	Menu Logout	Keluar dari halaman utama	Halaman Login Tampil dan aktif	Sesuai

## BAB V

### HASIL PEMBAHASAN

#### 5.1 Pembahasan Model

##### 1. Tampilan Pembuatan Model



Gambar 5.1 Pemodelan

Tampilan gambar diatas merupakan tampilan dalam pembuatan model pada metode C4.5, dalam tahap pembuatan model terdiri dari data penelitian, metode c45 itu sendiri, apply model, dan *performance*

##### 2. Tampilan Hasil Akurasi

The screenshot shows the 'PerformanceVector (Performance)' window with the 'Table / Plot View' selected. The table displays the following data:

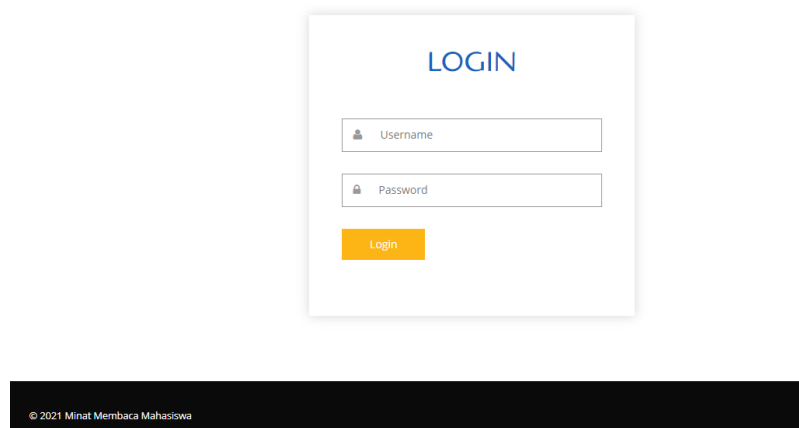
	true Sedang	true Tinggi	class precision
pred. Sedang	154	47	76.62%
pred. Tinggi	0	0	0.00%
class recall	100.00%	0.00%	

Gambar 5.2 Akurasi

Pada gambar 5.2 diatas merupakan tampilan hasil akurasi, yang mana hasil yang diperoleh dari data pada penelitian dengan menggunakan metode c4.5 sebesar 76,62%

## 5.2 Pembuatan Sistem

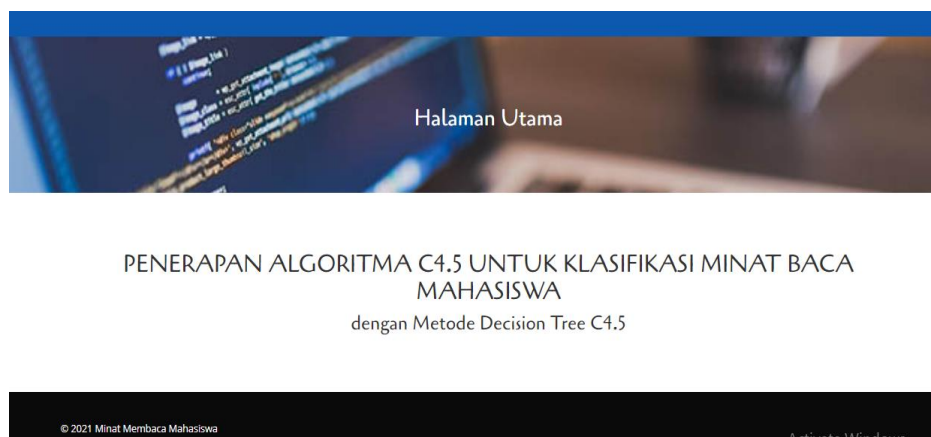
### 1. Tampilan Login Admin



**Gambar 5.3** Tampilan Login

Tampilan pada gambar 5.3 merupakan tampilan halaman login, dari tampilan diatas terdiri dari username dan password

### 2. Tampilan Halaman Admin



**Gambar 5.4** Tampilan Halaman Admin

Pada gambar 5.4 merupakan halaman admin, pada halaman ini terdiri dari beberapa menu yaitu menu Data User, menu Atribut, menu c45, menu hasil klasifikasi, serta menu logout

### 3. Tampilan Halaman Menu Data User

**Data User**

Nama

Username

Jenis Kelamin  
☐ Laki-laki  
☐ Perempuan

Usia

Jumlah data: 3

No	Nama	Jenis Kelamin	Usia	Username	
1	Coba Siswa	L	15	siswa	✖

Activate Windows

**Gambar 5.5** Tampilan Halaman User

Tampilan diatas merupakan tampilan halaman user, pada tampilan tersebut terdapat form pengisian data yang artinya sistem ini dapat menambah data dan menghapus data user

### 4. Tampilan Data Latih

**Data Latih**

Import data from excel  
 No file chosen

☒ Proses Mining

Jumlah data: 100

No	Nama	Jenis Kelamin	Usia	Jenis Bacaan	Tujuan Membaca	Lingkungan Baca	Minat Membaca
1	Fawwazadzka	P	Dewasa	non fiksi	menambah wawasan	teman	sedang
2	Wafida Mukrom	P	Dewasa	fiksi	menambah wawasan	teman	sedang

**Gambar 5.6** Tampilan Halaman Data Latih

Tampilan diatas merupakan tampilan halaman Data Latih, pada tampilan tersebut terdapat Tabel Data Latih

#### 5. Tampilan Hasil Klasifikasi



### Hasil Klasifikasi

Jumlah data: 1

No	Nama	Jenis Kelamin	Usia	Jenis Bacaan	Tujuan Membaca	Lingkungan Baca	Kelas Hasil	Id rule
1	Coba siswa	L	Dewasa	fiksi	menambah wawasan	Teman	sedang	1

**Gambar 5.7** Tampilan Hasil Klasifikasi

Tampilan diatas merupakan tampilan hasil klasifikasi, pada tampilan tersebut terdapat hasil klasifikasi minat baca mahasiswa sedang.

## **BAB VI**

### **KESIMPULAN DAN SARAN**

#### **6.1 Kesimpulan**

Berdasarkan hasil dari penelitian mengenai klasifikasi minat baca menggunakan Rapidminer dapat disimpulkan :

- A. Dapat diketahui hasil kinerja dan keefektifitasan dalam sistem klasifikasi dengan menggunakan metode C4.5, dengan mengukur akurasi sebesar 76.62%.
- B. Implementasi metode C4.5 dalam sistem klasifikasi minat baca mahasiswa dapat di implementasikan berdasarkan perhitungan  $CC = 4$  pada pengujian *white box*

#### **6.2 Saran**

Adapun saran berdasarkan penelitian ini berupa :

- 1. Dapat dilakukan penelitian lebih lanjut tentang metode klasifikasi minat baca mahasiswa, dalam pemeliharaan data karena semakin banyak data maka akan lebih baik
- 2. Penelitian selanjutnya dapat mengembangkan sistem menggunakan metode yang lain terkait tentang klasifikasi minat baca mahasiswa.



# CURRICULUM VITAE



Amran Pontoh, Dilahirkan di Manado pada tanggal 19 Februari 1996, Yang merupakan anak ke Tiga dari 3 bersaudara. Dari pasangan Bapak Pembantu Letnan Dua (Purn) TNI Muhidin Pontoh, dan Ibu almarhummah Suriani Ponamon, Kakak Pertama Nindy Elfira Pontoh,. Amd.Kep dan kakak ke dua Fikran Pontoh, S.M Dari desa Boroko, Kecamatan Kaidipang, Kabupaten Bolaang Mongondow Utara. Menjadi Mahasiswa Ichsan Gorontalo, Fakultas Ilmu Komputer Program Studi (S1) Teknik Informatika Angkatan 2014.

## **1. Riwayat Hidup**

### **a. Pendidikan Formal**

1. SDN 2 Bigo Pada Tahun 2002 – 2008.
2. MTS N 1 Kaidipang Pada Tahun 2008 – 2011.
3. SMKN 1 Kaidipang Pada Tahun 2011 – 2014.
4. Mahasiswa Pada Fakultas Ilmu Komputer Program Studi (S1) Teknik Informatika. Universitas Ichsan Gorontalo 2014.

### **b. Pendidikan Non Formal**

1. Peserta Orientasi Belajar Mahasiswa Baru (OBMB) di Universitas Ichsan Gorontalo Pada Tahun 2014.
2. Peserta KKLP Techopreneur Angkatan XXI Universitas Icsan Gorontalo di Desa Mutiara Kabupaten Boalemo Tahun 2018.

## DAFTAR PUSTAKA

- [1]. Gewati, M. (2016, Agustus 29). Retrieved from Kompas.Com:  
<http://edukasi.kompas.com/read/2016/08/29/07175131/minat.baca.indonesia.ada.di.urutan.ke-60.dunia>
- [2]. Prasetyo, E. (2014). *Data Mining : Mengolah Data Menjadi Informasi Menggunakan Matlab*. Gresik: Andi Offset.
- [3]. Saleh, A. (2015). Implmentasi Metode Klasifikasi Naive Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga.
- [4]. David Hartanto Kamagi dkk (2014). Implementasi Data Mining dengan Algoritma C4.5 untuk klasifikasi Tingkat Kelulusan Mahasiswa.
- [5]. Eki Ruziqa Maris (2014). Analisis Kepuasan Pelanggan Menggunakan Algoritma C4.5.
- [6]. Wisti Dwi Septiani (2014). Penerapan Algoritma C4.5 Untuk klasifikasi Penyakit Hepatitis.
- [7]. Fina nasari (2014). Penerapan Algoritma C4.5 Dalam Pemilihan Bidang Peminatan Program Studi Sistem Informasi Di Stimik Potensi Utama Medan.
- [8]. Ramadion, dan Listariono. (2009). *Peranan perpustakaan sekolah dalam upaya menumbuhkan minat baca siswa*. Malang: Universitas Negeri Malang
- [9]. Suyanto. (2014). *Artificial Intelligence*. Bandung: Informatika Bandung.
- [10]. Widodo, P. P., Handayanto, R. T., & Herlawati. (2013). *Penerapan Data Mining Dengan Matlab*. Bandung: Rekayasa Sains.

- [11]. Romario. (2013). *Penerapan Data Mining Pada Rsup Dr.Moh Hoesin Sumatera Selatan Untuk Pengelompokan Hasil Diagnosa Pasien Pengguna Asurans Kesehatan Miskin (askin)*. Palembang: Universitas Bina Darma Palembang.
- [12]. Gorunescu, F. (2011). *Data Mining Concept Model Technique*.
- [13].Jogiyanto HM,(2005).*Pengembangan sistem informasi yang berbasis komputer*.
- [14].Pressman, R. S. (2002). *Software Engineering: A Practitioners Approach*.  
Mc Graw-Hill.
- [15]. Alan Dennis, e. a. (2010). *System Analysis and Design with UML*. John Wiley and Sons.
- [16]. Sugiyono. (2009). *Metode Penelitian*. Bandung: CV. Alfabeta
- [17].*Perpustakaan Universitas Muhammadiyah Gorontalo. Catatan Pengunjung Perpustakaan UMGO*.

## 1. Lampiran Script(Program)

### 2. Login.php

```
<?php

session_start();

if ( isset($_SESSION['klasifikasi c45_id']) ) {

    header("location:index.php");

}

$login = 0;

if (isset($_GET['login'])) {

    $login = $_GET['login'];

}


if ($login == 1) {

    $komen = "Silahkan Login Ulang, Cek username dan Password Anda!!";

}


include_once "fungsi.php";

?>

<div class="content">

<!--login-->

<div class="login">

<div class="main-agileits">

<div class="form-w3agile">

<h3>PLEASE LOGIN</h3>

<?php

        if (isset($komen)) {

            display_error("Login failed");
```

```

    }

    ?>

<form action="cek-login.php" method="post">

<div class="key">

<i class="fa fa-user" aria-hidden="true"></i>

<input type="text" value="" name="username" required="" placeholder="Username">

<div class="clearfix"></div>

</div>

<div class="key">

<i class="fa fa-lock" aria-hidden="true"></i>

<input type="password" value="" name="password" required="" placeholder="Password">

<div class="clearfix"></div>

</div>

<input type="submit" value="Login">

</form>

</div>

</div>

</div>

<!--login-->

</div>

```

### 3. Koneksi.php

```

<?php
/*
 * Configuration Database
 */
return array(
    'host' => "localhost",
    'username' => "root",
    'password' => "",
    'dbname' => "klasifikasic45",
);

```

?>

#### 4. Metode C45 (Prosesmining.php)

```
<?php
function format_decimal($value){
    return round($value, 3);
}

//fungsi utama
function proses_DT($db_object, $parent, $kasus_cabang1, $kasus_cabang2) {
    echo "cabang 1<br>";
    pembentukan_tree($db_object, $parent, $kasus_cabang1);
    echo "cabang 2<br>";
    pembentukan_tree($db_object, $parent, $kasus_cabang2);
}

//fungsi proses dalam suatu kasus data
function pembentukan_tree($db_object, $N_parent, $kasus) {
    //mengisi kondisi
    if ($N_parent != "") {
        $kondisi = $N_parent . " AND " . $kasus;
    } else {
        $kondisi = $kasus;
    }
    echo $kondisi . "<br>";
    //cek data heterogen / homogen???
    $cek = cek_heterohomogen($db_object, 'kelas_asli', $kondisi);
    if ($cek == 'homogen') {
        echo "<br>LEAF ||";
        $sql_keputusan = $db_object->db_query("SELECT DISTINCT(kelas_asli) FROM "
            . "data_latih WHERE $kondisi");
        $row_keputusan = $db_object->db_fetch_array($sql_keputusan);
        $keputusan = $row_keputusan[0];
        //insert atau lakukan pemangkasan cabang
        pangkas($db_object, $N_parent, $kasus, $keputusan);
    } //jika data masih heterogen
    else if ($cek == 'heterogen') {

        if ($kondisi != "") {
            $kondisi_kelas_asli = $kondisi . " AND ";
        }
        $jml_sedang = jumlah_data($db_object, "$kondisi_kelas_asli kelas_asli='sedang'");
        $jml_tinggi = jumlah_data($db_object, "$kondisi_kelas_asli kelas_asli='tinggi'");
```

```

$jml_total = $jml_sedang + $jml_tinggi;
echo "Jumlah data = " . $jml_total . "<br>";
echo "Jumlah Sedang = " . $jml_sedang . "<br>";
echo "Jumlah Tinggi = " . $jml_tinggi . "<br>";

//hitung entropy semua
$entropy_all = hitung_entropy($jml_sedang, $jml_tinggi);
echo "Entropy All = " . $entropy_all . "<br>";

$nilai_usia = array();
$nilai_usia = cek_nilaiAtribut($db_object, 'usia', $kondisi);
$jmlUsia = count($nilai_usia);

echo "<table class='table table-bordered table-striped table-hover'>";
echo "<tr><th>Nilai Atribut</th><th>Jumlah data</th><th>Jumlah Sedang</th><th>Jumlah Tinggi</th> ". "
<th>Entropy</th><th>Gain</th><tr>";

$db_object->db_query("TRUNCATE gain");
//hitung gain atribut KATEGORIKAL
hitung_gain($db_object, $kondisi, "jenis_kelamin", $entropy_all, "jenis_kelamin='L'", "jenis_kelamin='P'", "",
"", "");

//hitung gain atribut KATEGORIKAL
if($jmlUsia!=1){
    $NA1Usia="usia='$nilai_usia[0]'";
    $NA2Usia="";
    $NA3Usia="";
    if($jmlUsia==2){
        $NA2Usia="usia='$nilai_usia[1]'";
    }else if ($jmlUsia==3){
        $NA2Usia="usia='$nilai_usia[1]'";
        $NA3Usia="usia='$nilai_usia[2]'";
    }
    hitung_gain($db_object, $kondisi, "usia", $entropy_all, $NA1Usia, $NA2Usia, $NA3Usia, "", "");
}

hitung_gain($db_object, $kondisi, "jenisbacaan", $entropy_all, "jenisbacaan='fiksi'", "jenisbacaan='non fiksi'",
"", "", "");

hitung_gain($db_object, $kondisi, "tujuanmembaca", $entropy_all, "tujuanmembaca='Menambah
Wawasan'", "tujuanmembaca='Motivasi'", "", "", "");

hitung_gain($db_object, $kondisi, "lingkunganbaca", $entropy_all, "lingkunganbaca='Teman'",
"lingkunganbaca='Asrama'", "", "", "");

//ambil nilai gain terbesar
$sql_max = $db_object->db_query("SELECT MAX(gain) FROM gain");

```

```

$row_max = $db_object->db_fetch_array($sql_max);
$max_gain = $row_max[0];
$sql = $db_object->db_query("SELECT * FROM gain WHERE gain=$max_gain");
$row = $db_object->db_fetch_array($sql);
$atribut = $row[2];
echo "Atribut terpilih = " . $atribut . ", dengan nilai gain = " . $max_gain . "<br>";
echo "<br>=====<br>";

//jika max gain = 0 perhitungan dihentikan dan mengambil keputusan
if ($max_gain == 0) {
    echo "<br>LEAF ";
    $Nsedang = $kondisi . " AND kelas_asli='Sedang'";
    $Ntinggi = $kondisi . " AND kelas_asli='tinggi'";

    $jumlahsedang = jumlah_data($db_object, "$Nsedang");
    $jumlahtinggi = jumlah_data($db_object, "$Ntinggi");

    if($jumlahsedang >= $jumlahtinggi) {
        $keputusan = 'sedang';
    }

    else {
        $keputusan = 'tinggi';
    }
    //insert atau lakukan pemangkasan cabang
    pangkas($db_object, $N_parent, $kasus, $keputusan);
}
//jika max_gain >0 lanjut..
else {
    //jenis kelamin terpilih
    if ($atribut == "jenis_kelamin") {
        proses_DT($db_object, $kondisi, "($atribut='L')", "($atribut='P')");
    }

    //usia terpilih
    if ($atribut == "usia") {
        //jika nilai atribut 3
        if($jmlUsia==3){
            //hitung rasio
            $cabang = array();
            $cabang = hitung_rasio($db_object, $kondisi,
'usia',$max_gain,$nilai_usia[0],$nilai_usia[1],$nilai_usia[2],",");
            $exp_cabang = explode(" ", $cabang[1]);
            proses_DT($db_object, $kondisi, "($atribut='$cabang[0]')", "($atribut='$exp_cabang[0]' OR
$atribut='$exp_cabang[1]')");
        }
        //jika nilai atribut 2
        else if($jmlUsia==2){

```



```

        proses_DT($db_object, $kondisi , "($atribut='$_nilai_usia[0]')", "($atribut='$_nilai_usia[1]')");
    }
}

```

#### 4. Klasifikasi.php

```

<div class="typo-w3">
<div class="container">
<h2 class="tittle">Klasifikasi</h2>
<?php
    //object database class
    $db_object = new database();
    $sql = "SELECT * FROM data_soal";
    $query = $db_object->db_query($sql);
    $jumlah = $db_object->db_num_rows($query);

    $pesan_error = $pesan_success = "";
    if (isset($_GET['pesan_error'])) {
        $pesan_error = $_GET['pesan_error'];
    }
    if (isset($_GET['pesan_success'])) {
        $pesan_success = $_GET['pesan_success'];
    }

    if (isset($_POST['submit'])) {
        $success = true;
        $lihat_hasil = false;
        $pesan_gagal = $pesan_sukses = "";
        $idSiswa = $_SESSION['klasifikasi c45_id_siswa'];
        if ($idSiswa <= 0) {
            $success = false;
            $pesan_gagal = "Anda bukan siswa";
        }

        if (sudah_klasifikasi($db_object, $idSiswa)) {
            $success = false;
            $lihat_hasil = true;
            $pesan_gagal = "Anda sudah melakukan klasifikasi";
        }

        if ($success) {
            $val_in = $di_jawab_jenisbacaan = $di_jawab_tujuanmembaca = $di_jawab_lingkunganbaca = array();
            foreach ($_POST['soal'] as $key => $value) {
                if (empty($value)) {
                    $success = false;
                    $pesan_gagal = "Ada yang belum diisi";
                }
            }
        }
    }
}

```

```

        break;
    }
    //key = id_soal, value=jawaban A/B/C/D
    $val_in[] = "(" . $_SESSION['klasifikasi c45_id'] . "," . $idSiswa .
        "," . $key . "," . $value . ")";
    if ($value == 'A') {
        $di_jawab_jenisbacaan[] = $key;
    }
    if ($value == 'B') {
        $di_jawab_tujuanmembaca[] = $key;
    }
    if ($value == 'C') {
        $di_jawab_lingkunganbaca[] = $key;
    }
}

//insert ke jawaban_kuisisioner
if ($idSiswa > 0) {
    $value_sql_to_in = implode(",", $val_in);
    $sql_in_jawaban = "INSERT INTO jawaban_pilihanbaca
        (id_user, id_siswa, id_soal, jawaban)
        VALUES " . $value_sql_to_in;
    $db_object->db_query($sql_in_jawaban);

    //hitung naive bayes
    $siswa = get_data_siswa($db_object, $idSiswa);
    $jenisbacaan = count($di_jawab_jenisbacaan);
    $tujuanmembaca = count($di_jawab_tujuanmembaca);
    $lingkunganbaca = count($di_jawab_lingkunganbaca);

    $hasil = klasifikasi($db_object, $siswa['jenis_kelamin'], $siswa['usia'], $jawaban_jenisbacaan,
        $jawaban_tujuanmembaca, $jawaban_lingkunganba

```



## UPT PERPUSTAKAAN

(NPP. 7501102D2000001, Sertifikat Akreditasi "A" No: 0191/LAP.PT/XI.2020)

### UNIVERSITAS MUHAMMADIYAH GORONTALO

Jl. Prof. Dr. Mansoer Pateda, Desa Pentadio Timur, Kec. Telaga Biru, Kab. Gorontalo, Kode Pos 96181  
Laman: lib.umgo.ac.id, Surel: perpus@umgo.ac.id, HP/WA. 0822 9335 9959

#### REKOMENDASI

Nomor 004/A/PERPUST-UMGo/III/2021

Berdasarkan Surat Lembaga Penelitian Universitas Ichsan Gorontalo Nomor 1252/PIP/LEMLIT-UNISA/GTO/III/2021 tanggal 15 Maret 2021 tentang Permohonan Izin Penelitian, dengan ini kami memberikan Rekomendasi kepada:

Nama	: Amran Pontoh
NIM	: T3114047
Jenis Kelamin	: Laki-laki
Pekerjaan	: Mahasiswa
Alamat	: Desa Boroko Timur Kec. Kaidipang Kab. Bolaang Mongondow Utara
Judul Penelitian	: Penerapan Algoritma C4.5 Untuk Klasifikasi Minat Baca Mahasiswa
Lokasi Penelitian	: Perpustakaan Universitas Muhammadiyah Gorontalo

Dalam melakukan penelitian agar memperhatikan ketentuan sebagai berikut:

1. Menjaga keamanan dan ketertiban
2. Peneliti wajib menyampaikan hasil penelitiannya kepada Perpustakaan Universitas Muhammadiyah Gorontalo

Demikian Surat Rekomendasi ini diberikan untuk dipergunakan sebagaimana mestinya

*Nasrun minallah wa fathun qarieb,  
Wassalamu'alaikum Warohmatullahi Wabarukatu*

Perpustakaan UMGo  
Rektor  
  
Dr. H. Saida Gani, M.Sos.I  
NPM 1183190

Tembusan:

1. Yth, Rektor UMGo (sebagai laporan)
2. Arsip



**KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN  
UNIVERSITAS ICHSAN  
(UNISAN) GORONTALO**

SURAT KEPUTUSAN MENDIKNAS RI NOMOR 84/D/O/2001

Jl. Achmad Nadjamuddin No. 17 Telp (0435) 829975 Fax (0435) 829976 Gorontalo

**SURAT REKOMENDASI BEBAS PLAGIASI**

No. 0728/UNISAN-G/S-BP/VI/2021

Yang bertanda tangan di bawah ini:

Nama : Sunarto Taliki, M.Kom  
NIDN : 0906058301  
Unit Kerja : Pustikom, Universitas Ichsan Gorontalo

Dengan ini Menyatakan bahwa :

Nama Mahasiswa : AMRAN PONTOH  
NIM : T3114047  
Program Studi : Teknik Informatika (S1)  
Fakultas : Fakultas Ilmu Komputer  
Judul Skripsi : PENERAPAN ALGORITMA C4.5  
KLASIFIKASI MINAT BACA MAHASISWA UNTUK

Sesuai dengan hasil pengecekan tingkat kemiripan skripsi melalui aplikasi Turnitin untuk judul skripsi di atas diperoleh hasil Similarity sebesar 27%, berdasarkan SK Rektor No. 237/UNISAN-G/SK/IX/2019 tentang Panduan Pencegahan dan Penanggulangan Plagiarisme, bahwa batas kemiripan skripsi maksimal 35% dan sesuai dengan Surat Pernyataan dari kedua Pembimbing yang bersangkutan menyatakan bahwa isi softcopy skripsi yang diolah di Turnitin SAMA ISINYA dengan Skripsi Aslinya serta format penulisannya sudah sesuai dengan Buku Panduan Penulisan Skripsi, untuk itu skripsi tersebut di atas dinyatakan BEBAS PLAGIASI dan layak untuk diujikan.

Demikian surat rekomendasi ini dibuat untuk digunakan sebagaimana mestinya.

Gorontalo, 01 Juni 2021

Tim Verifikasi,



Sunarto Taliki, M.Kom

NIDN. 0906058301

Tembusan :

1. Dekan
2. Ketua Program Studi
3. Pembimbing I dan Pembimbing II
4. Yang bersangkutan
5. Arsip

T3114047 AMRAN PONTOK

## PENERAPAN ALGORITMA C4.5 UNTUK KLASIFIKASI MINAT BA...

Sources Overview

27%

OVERALL SIMILARITY

1	www.scribd.com	7%
2	etheses.uin-malang.ac.id	4%
3	sir.stikom.edu	2%
4	digilib.binadarma.ac.id	2%
5	id.123dok.com	2%
6	docplayer.info	1%
7	123 dok.com	<1%
8	repository.uinjkt.ac.id	<1%
9	ejurnal.stmik-budidarma.ac.id	<1%
10	journal.universitaspahlawan.ac.id	<1%
11	www.kom-unisan.ac.id	<1%
12	eprints.unisnu.ac.id	<1%
13	eprints.dinus.ac.id	<1%
14	repository.unikom.ac.id	<1%
15	nurjayanti.blog.widyatama.ac.id	<1%
16	indraharmukti.blogspot.com	<1%

17	www.integraphone.fr	
	INTERNET	<1%
18	sistemasi.ftik.unisi.ac.id	
	INTERNET	<1%
19	irin-halid.blogspot.com	
	INTERNET	<1%
20	edrianhadinata.wordpress.com	
	INTERNET	<1%
21	fhiezasetia102513.blogspot.com	
	INTERNET	<1%
22	Erлита Mentari Lahengking, Inggriani Elim, Rudy Pusung. "ANALISIS PIUTANG USAHA PADA PERUSAHAAN DAERAH AIR MINUM (PDA...	
	CROSSREF	<1%

Excluded search repositories:

- Submitted Works

Excluded from Similarity Report:

- Small Matches (less than 25 words).

Excluded sources:

-