**LAPORAN PENELITIAN**

# PENERAPAN ALGORITMA C4.5 UNTUK MEMPREDIKSI KELULUSAN MAHASISWA 2020FAKULTAS SAINS DAN TEKNOLOGI (STUDI KASUS: TEKNOLOGI INFORMASI, TEKNIK LINGKUNGAN, ARSITEKTUR,KIMIA, BIOLOGI)


**Peneliti:**

**ANGGUN MAY ERDELITA**

NIM.190705041


| | |
|---|---|
| **Jenis Penelitian** | **Terapan** |
| **Bidang Ilmu Kajian** | **Sistem Pendukung Keputusan, Sistem Pakar** |
| **Dosen Peneliti** | **Khairan AR, M.Kom** |


**UNIVERSITAS ISLAM NEGERI AR-RANIRY BANDA ACEH**

**FAKULTAS SAINS DAN TEKNOLOGI**

**PRODI TEKNOLOGI INFORMASI**


**APRIL 2023**

**LEMBARAN IDENTITAS DAN PENGESAHAN LAPORAN PENELITIANPUSAT PENELITIAN DAN PENERBITAN LP2M UIN AR-RANIRY TAHUN 2023**

1. a. Judul : Penerapan Algotitma C45 Untuk Memprediksi Kelulusan Mahasiswa 2020 Fakultas Sains dan Teknologi **(Studi Kasus: Teknologi Informasi, Teknik Lingkungan, Arsitektur, Kimia, Biologi)**

   b. Jenis Penelitian : Penelitian Terapan
   c. No. Registrasi : -
   d. Bidang Ilmu yang diteliti : Sistem Pendukung Keputusan, Sistem Pakar

2. Peneliti
   a. Nama Lengkap : Anggun May Erdelita
   b. Jenis Kelamin : Perempuan
   c. NIM : 190705041
   h. Fakultas/Prodi : Sains dan Teknologi / Teknologi Informasi

   i. Anggota Peneliti 1
      Nama Lengkap : Khairan AR, M.Kom
      Jenis Kelamin : Laki laki
      Fakultas/Prodi : Sains dan Teknilogi/ Teknologi Informasi

   j. Anggota Peneliti 2 *(Jika Ada)*
      Nama Lengkap : Hendri Ahmadian, S.Si, M.I.M
      Jenis Kelamin : Laki Laki
      Fakultas/Prodi : Sains dan Teknilogi/ Teknologi Informasi

3. Lokasi Kegiatan : Fakultas Sains dan Teknologi UIN Ar-Raniry Banda Aceh
4. Jangka Waktu Pelaksanaan : 6 (Enam) Bulan
5. Tahun Pelaksanaan : 2023
6. Jumlah Anggaran Biaya : -
7. Sumber Dana : Mandiri
8. *Output* dan *Outcome* : -

Mengetahui,
Dosen Pembimbing I

**Khairan AR, M.Kom**
NIP. 198607042014031001

Banda Aceh, 2 April 20223Pelaksana,

**Anggun May Erdelita**
NIM. 19705041

Menyetujui:

Ketua Prodi. Teknologi Informasi

**Ima Dwitawati**

NIP. 198210132014032002

# ABSTRAK


| | |
|---|---|
| Nama | : Anggun May Erdelita |
| NIM | 190705041 |
| Program Studi | : Teknologi Informasi |

| | |
|---|---|
| Judul | : Penerapan Algoritma C4.5 Untuk Memprediksi Kelulusan Mahasiswa Fakultas Sains dan Teknologi (Studi Kasus: Teknologi Informasi, Teknik Lingkungan, Arsitektur, Kimia, Biologi) |
| Tanggal Sidang | : 06 April 2023 |
| Jumlah Halaman | :103 Halaman |
| Pembimbing I | : Khairan AR, M.Kom |
| Pembimbing II | : Hendri Ahmadian, S.Si.,M.I.M |
| Kata Kunci | : Algoritma C4.5, Prediksi, *Confusion Matrix* |


Universitas Islam Negeri Ar-Raniry Banda Aceh merupakan salah satu Perguruan Tinggi Negeri di Indonesia yang memiliki 9 fakultas, salah satunya Fakultas Sains dan Teknologi. Setiap tahunnya, kuota mahasiswa yang diterima semakin bertambah, namun tidak semua mahasiswa dapat lulus tepat waktu sesuai dengan masa studi yang ditempuh. Sehingga, mengakibatkan penumpukan jumlah mahasiswa yang tidak lulus sesuai dengan masa periode kelulusannya. Berdasarkan latar belakang pemasalahan, penelitian ini bertujuan untuk mengetahui hasil prediksi kelulusan mahasiswa Fakultas Sains dan Teknologi UIN Ar-Raniry Banda Aceh pada lima program studi yaitu Arsitektur, Teknik Lingkungan Biologi, Kimia, dan Teknologi Informasi dengan menggunakan analisa deskriptif kuantitatif. Metode yang digunakan adalah Algoritma C4.5 dengan metode penarikan sampel menggunakan teknik *Stratified Random Sampling*. Penggunaan bahasa pemrograman PHP dan *framework Codeigniter* pada sistem prediksi kelulusan. Jumlah data yang yang digunakan sebanyak 437 meliputi 306 data *training*, dan 131 data *testing*. Hasil dari penelitian menunjukkan bahwa atribut IPK paling berpengaruh dalam memprediksi kelulusan dengan tiga keluaran yaitu lulus lebih awal, tepat waktu, dan terlambat. Dari hasil prediksi memperlihatkan bahwa program studi Teknologi Informasi memiliki kelulusan terlambat dan lebih awal paling banyak diantara 4 program studi lainnya. Akurasi yang diperoleh sebesar 88%, *presisi* 96%, dan *recall* 94% dengan klasifikasi baik yang membuktikan bahwa Algoritma C4.5 dapat digunakan untuk melakukan prediksi.

# DAFTAR ISI

**ABSTRAK** ..... i
**DAFTAR ISI** ..... ii
**DAFTAR GAMBAR** ..... v
**DAFTAR TABEL** ..... viii
**DAFTAR LAMPIRAN** ..... ix
**BAB I** ..... 1
**PENDAHULUAN** ..... 1
**I.1 Latar Belakang** ..... 1
**I.2 Rumusan Masalah** ..... 3
**I.3 Tujuan Penelitian** ..... 3
**I.4 Batasan Masalah** ..... 3
**I.5 Manfaat Penelitian** ..... 4
**I.6 Sistematika Penulisan** ..... 5
**BAB II** ..... 6
**KAJIAN PUSTAKA** ..... 6
**II.1 Penelitian Terdahulu** ..... 6
**II.2 Kajian Teoritis** ..... 8
II.2.1 Data Mining ..... 8
II.2.2 Fungsi Data Mining ..... 8
II.2.3 Ruang Lingkup Data Mining ..... 9
II.2.4 Proses Data Mining ..... 9
II.2 .5 Penerapan Data Mining ..... 10
II.3 *Decision Tree* (Pohon Keputusan) ..... 10
II.4 **Algoritma C4.5** ..... 12
II.5 ***Classification* (Klasifikasi)** ..... 12
II.6 ***Confusion Matrix*** ..... 13
II.7 ***Website*** ..... 15
II.8 **Basis Data** ..... 15
II.9 ***MySQL*** ..... 16
II.10 ***PhpMyadmin*** ..... 16

II.11 ***Codeigniter*** ..... 16
II.12 ***PHP*** ..... 18
II.13 ***UML* (*Unifield Modelling Languange*)** ..... 18
II.14 ***Framework*** ..... 18
II.15 ***Bootstrap*** ..... 19
II.16 ***XAMPP*** ..... 19
II.17 **Kerangka Penelitian** ..... 21
**BAB III** ..... 22
**METODE PENGEMBANGAN PRODUK** ..... 22
**III.1 Metodologi Pengembangan Produk** ..... 22
III.1.1 Algoritma C4.5 ..... 24
**III.2 Waktu dan Tempat Penelitian** ..... 26
**III.3 Perangkat Penelitian** ..... 26
**III.4 Deskripsi Sistem** ..... 27
**III.5 Perancangan Sistem** ..... 31
III.5.1 Desain *Input* ..... 31
III.5.2 Desain *Output* ..... 32
III.5.3 Desain Proses ..... 33
III.5.4 *Use Case* Diagram ..... 34
III.5.5 *Activity Diagram* ..... 38
III.5.6 *Sequence Diagram* ..... 47
III.5.7 *Class Diagram* ..... 62
**III.6 Populasi** ..... 62
**III.7 Sampel** ..... 63
**III.8 Teknik Pengambilan Sampel** ..... 64
**III.9 Jenis dan Sumber Data** ..... 65
**III.10 Teknik Pengumpulan Data** ..... 65
**III.11 Teknik Pengolahan Data** ..... 66
**III.12 Teknik Analisis** ..... 66
III.12.1 Teknik Analisis Data ..... 66
III.12.2 Teknik Analisis Sistem ..... 68
**III.13 Teknik Pengujian** ..... 68
III.13.1 Teknik Pengujian Data ..... 68
III.13.2 Teknik Pengujian Sistem ..... 69

**BAB IV** ..... 70
**HASIL DAN PEMBAHASAN** ..... 70
**IV.1 Perancangan Tampilan Antarmuka (*Interface*)** ..... 70
**IV.2 Implementasi Sistem** ..... 86
**IV.3 Dataset** ..... 95
**IV.4 *Confusion Matrix*** ..... 96
**BAB V** ..... 99
**KESIMPULAN DAN SARAN** ..... 99
**V.1 Kesimpulan** ..... 99
**V.2 Saran** ..... 99
**DAFTAR PUSTAKA** ..... 100
**LAMPIRAN** ..... 103

# DAFTAR GAMBAR

**Gambar 3. 1** Tahapan Penelitian.....22
**Gambar 3. 2** Diagram Blok Sistem.....24
**Gambar 3. 3** Flowchart Penentuan Root.....28
**Gambar 3. 4** Flowchart Penentuan Cabang .....29
**Gambar 3. 5** Flowchart Penentuan Node.....30
**Gambar 3. 6** Use Case Login.....34
**Gambar 3. 7** Use Case Data Mahasiswa.....35
**Gambar 3. 8** Use Case Prediksi .....36
**Gambar 3. 9** Use Case Laporan Hasil.....37
**Gambar 3. 10** Activity Admin Login.....38
**Gambar 3. 11** Activity User Login .....39
**Gambar 3. 12** Activity Admin pada data mahasiswa.....40
**Gambar 3. 13** Activity User pada Data Mahasiswa.....41
**Gambar 3. 14** Activity Prediksi .....42
**Gambar 3. 15** Activity Laporan Admin.....43
**Gambar 3. 16** Activity Laporan User .....44
**Gambar 3. 17** Activity Ubah password .....45
**Gambar 3. 18** Activity Logout.....46
**Gambar 3. 19** Sequence Login .....47
**Gambar 3. 20** Sequence Admin Tambah Data Mahasiswa .....48
**Gambar 3. 21** Sequence Admin Ubah Data Mahasiswa.....49
**Gambar 3. 22** Sequence Admin Hapus Data Mahasiswa .....50
**Gambar 3. 23** Sequence Lihat Data Mahasiswa .....51
**Gambar 3. 24** Sequence Search Data Mahasiswa.....52
**Gambar 3. 25** Sequence Admin Input Data Training .....53
**Gambar 3. 26** Sequence Admin Import Data Training.....54
**Gambar 3. 27** Sequence Rule .....55
**Gambar 3. 28** Sequence Input Target .....56
**Gambar 3. 29** Sequence Import Data Multitarget.....57
**Gambar 3. 30** Sequence Lihat Laporan Hasil Prediksi.....58

**Gambar 3. 31** Sequence Cetak Laporan .....59
**Gambar 3. 32** Sequence Ubah Password.....60
**Gambar 3. 33** Sequence Logout .....61
**Gambar 3. 34** Class Diagram.....62
**Gambar 3. 35** Black box testing .....69
**Gambar 4. 1** Halaman Login ..... 70
**Gambar 4. 2** Halaman Dashboard Admin .....71
**Gambar 4. 3** Halaman Dashboard User.....72
**Gambar 4. 4** Halaman Data Mahasiswa .....72
**Gambar 4. 5** Halaman Tambah Data Mahasiswa .....73
**Gambar 4. 6** Halaman Update Data Mahasiswa.....74
**Gambar 4. 7** Halaman Hapus Data Mahasiswa .....75
**Gambar 4. 8** Halaman Data Mining .....75
**Gambar 4. 9** Halaman Tambah Data Training .....76
**Gambar 4. 11** Halaman Ubah Data Training.....77
**Gambar 4. 12** Halaman Hapus Data Traning .....78
**Gambar 4. 13** Halaman Data User.....79
**Gambar 4. 14** Halaman Admin Tambah Data User.....80
**Gambar 4. 15** Halaman Rules.....81
**Gambar 4. 16** Halaman Entropy dan Gain .....81
**Gambar 4. 17** Halaman Proses Prediksi Kelulusan .....82
**Gambar 4. 18** Halaman Import Data Training.....83
**Gambar 4. 19** Halaman Import Data Testing .....83
**Gambar 4. 20** Halaman Cetak Laporan .....84
**Gambar 4. 21** Halaman Ubah Password.....85
**Gambar 4. 22** Halaman Login .....86
**Gambar 4. 23** Halaman Dashboard .....87
**Gambar 4. 24** Data Training.....88
**Gambar 4. 25** Input Data Training.....88
**Gambar 4. 26** Import Data Training .....89
**Gambar 4. 27** Data Testing.....89
**Gambar 4. 28** Input Data Testing .....90

**Gambar 4. 29** Import Data Testing .....90
**Gambar 4. 30** Pohon Keputusan .....91
**Gambar 4. 31** Perhitungan Entropy & Gain .....91
**Gambar 4. 32** Proses Prediksi .....92
**Gambar 4. 33** Proses Akurasi Data .....92
**Gambar 4. 34** Cetak Laporan Prediksi .....93
**Gambar 4. 35** Cetak Laporan Akurasi Prediksi .....93
**Gambar 4. 36** Halaman User .....94
**Gambar 4. 37** Input Data User .....94
**Gambar 4. 38** Hasil Prediksi .....98

# DAFTAR TABEL

**Tabel 3. 1**. Perangkat penelitian .....26
**Tabel 3. 2** Desain Input .....31
**Tabel 3. 3** Desain Output.....32
**Tabel 3. 4** Desain Proses .....33
**Tabel 3. 5** Populasi Angkatan 2020.....63
**Tabel 3. 6** Jumlah Sampel Mahasiswa Angkatan 2020 .....65
**Tabel 3. 7** Predikat IP Program Studi S1 .....67
**Tabel 3. 8** Predikat IPK Program Studi S1 .....67
**Tabel 3. 9** Keterangan perkuliahan.....67

# DAFTAR LAMPIRAN

Lampiran I : Lampiran Data

Lampiran II : Lampiran Program Code

# BAB I

# PENDAHULUAN

## I.1 Latar Belakang

Dewasa ini, teknologi informasi mengalami perkembangan yang sangat pesat. Kebutuhan akan sebuah sistem yang terkomputerisasi dapat mencakup berbagai bidang. Salah satunya bidang pendidikan. Setiap perguruan tinggi baik swasta maupun negeri sangat membutuhkan sistem komputerisasi yang akurat, relevan, efisien, dan cepat. Mahasiswa merupakan indikator penting di sebuah instansi perguruan tinggi. Salah satu penentu kualitas perguruan tinggi adalah jumlah kemampuan mahasiswa yang dapat menyelesaikan masa studi tepat waktu. Akan tetapi, terdapat beberapa faktor yang menghambat mahasiswa untuk dapat lulus tepat waktu baik faktor secara internal maupun faktor eksternal.

Universitas Islam Negeri Ar-Raniry (UIN Ar-Raniry) adalah salah satu Perguruan Tinggi Islam Negeri (PTKIN) yang terletak di Banda Aceh dengan nama sebelumnya Institut Agama Islam Ar-Raniry (IAIN). Tepat pada 05 Oktober 2013, perguruan tinggi tersebut merubah nama dari Institut menjadi Universitas melalui PERPRES No. 64 Tahun 2013. Saat ini UIN Ar-Raniry telah memiliki 9 fakultas dengan program studi yang terdiri dari 2 program D3, 42 Program S1 dan 9 Program Pascasarjana. Salah satu fakultas yang terdapat di UIN Ar-Raniry yaitu Fakultas Sains dan Teknologi yang telah memiliki 6 program studi, antara lain Teknologi Informasi, Teknik Lingkungan, Kimia, Fisika, Biologi, dan Arsitektur.

Berdasarkan data alumni pada *website* resmi Fakultas Sains dan Teknologi, mahasiswa yang lulus per Agustus 2022 sebanyak 166 orang. Hal ini tidak sebanding dengan jumlah mahasiswa baru yang masuk pada program studi Fakultas Sains dan Teknologi. Tentunya hal ini sangat mempengaruhi penilaian akreditasi sehingga perlu untuk mengetahui parameter apa saja yang mempengaruhi mahasiswa dapat menyelesaikan studi. Oleh karena itu, perlu adanya suatu sistem yang dapat digunakan untuk memprediksi kelulusan mahasiswa Angkatan 2020.

Penelitian mengenai prediksi kelulusan telah banyak dilakukan seperti penelitian oleh (Anwar F.F,dkk, 2022) dengan 16 atribut dan 1 atribut target yang diuji menggunakan 75 data *testing*. Diperoleh hasil 69 data mahasiswa diprediksi benar, dan 6 data mahasiswa diprediksi salah dengan hasil nilai *accuracy* 92%, *error rate* 8%, nilai *AUC* sebesar 0,7138. Penelitian lainnya yang dilakukan Susi Mashlahah yang berjudul "Prediksi Kelulusan Mahasiswa Menggunakan Metode *Decision Tree* Dengan Penerapan Algoritma C4.5" memperoleh hasil akurasi kecocokan pada sistem mencapai 82,79% dengan atribut asal daerah, jenis sekolah, jalur masuk, pengalaman pesantren, IPK, dan IP semester 1 sampai 5. Prediksi kelulusan lain yang menggunakan algoritma C4.5 diteliti oleh Lydia Yohana Lumban Gaol, dkk, 2021 terhadap mahasiswa Stikom Tunas Bangsa Prodi Sistem Informasi memperoleh *accuracy* 90.00%, *precision* 91.38% dan *recall* 98.15%. Hal ini menunjukkan atribut IPK memiliki pengaruh paling tinggi dalam menentukan kelulusan mahasiswa.

Berdasarkan pemaparan diatas dibangun sebuah sistem informasi prediksi kelulusan mahasiswa Fakultas Sains dan Teknologi menggunakan bahasa pemrograman PHP dan *framework Codeignitier* dengan judul "**Penerapan Algoritma C4.5 Untuk Memprediksi Kelulusan Mahasiswa 2020 Fakultas Sains dan Teknologi (Studi Kasus: Teknologi Informasi, Teknik Lingkungan, Arsitektur, Kimia, Biologi)"**. Metode yang digunakan dalam penelitian ini adalah klasifikasi. Model klasifikasi dibuat untuk mengidentifikasi pola data untuk kelas status lulus lebih awal, lulus tepat waktu, dan lulus terlambat dari hasil penentuan pola training data.

Algoritma yang diterapkan pada penelitian ini adalah Algoritma C4.5 karena memiliki kelebihan untuk dapat menampilkan sebuah pohon keputusan dalam bentuk *rule* (aturan) dan mengklasifikasikan suatu kejadian. Tujuan dilakukannya penelitian ini untuk memprediksi kelulusan mahasiswa pada 5 program studi yang terdapat di Fakultas Sains dan Teknologi Angkatan 2020. Pemilihan data angkatan 2020 sebagai data uji didasari untuk sebagai acuan data pertama yang akan diprediksi dengan menggunakan data *training* angkatan sebelumnya dengan menggunakan tujuh atribut yaitu IP semester 1-5, IPK, dan keterangan perkuliahan.

Untuk pengukuran kinerja dilakukan menggunakan *Confusion Matrix* dengan tujuan memperoleh nilai *accuracy, precision,* dan *recall.* Data *training* yang akan digunakan oleh peneliti adalah data mahasiswa dari 5 Program Studi Fakultas Sains dan Teknologi Angkatan 2020. Dari hasil pengolahan data tersebut akan diketahui informasi kelulusan mahasiswa yang dapat dijadikan dasar dalam pengambilan informasi demi meningkatkan jumlah lulusan.

## I.2 Rumusan Masalah

Berdasarkan latar belakang diatas, diperoleh rumusan masalah sebagai berikut :

1. Bagaimana merancang sebuah sistem yang dapat memprediksi kelulusan mahasiswa Fakultas Sains dan Teknologi menggunakan algoritma C4.5 ?
2. Bagaimana tingkat akurasi dari hasil analisa prediksi kelulusan mahasiswa dengan menggunakan metode *Confusion Matrix* ?

## I.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah membuat sistem informasi yang dapat memprediksi kelulusan lebih awal, tepat waktu, dan terlambat serta mengetahui tingkat akurasi dari hasil prediksi kelulusan mahasiswa Fakultas Sains dan Teknologi Angkatan 2020.

## I.4 Batasan Masalah

Batasan masalah pada penelitian ini diantaranya sebagai berikut :

1. Ruang lingkup penelitian dibatasi pada data 5 program studi Fakultas Sains dan Teknologi yaitu Teknologi Informasi, Teknik Lingkungan, Arsitek, Kimia, dan Biologi.
2. Sistem prediksi kelulusan berbasis web menggunakan bahasa pemrograman PHP dan *framework Codeignitier.*
3. Atribut yang digunakan dalam perhitungan C4.5 yaitu IP semester 1-5, IPK, dan keterangan perkuliahan.
4. Menggunakan *Sublime Text Editor*.
5. Data yang digunakan adalah data mahasiswa dari 5 program studi yang diperoleh dari Sistem Informasi Akademik (SIAKAD).

## I.5 Manfaat Penelitian

Adapun manfaat dari penelitian ini, antara lain :

a. Manfaat Teoritis
   1. Mampu mengetahui proses penyelesaian prediksi kelulusan mahasiswa Fakultas Sains dan Teknologi UIN Ar-Raniry Banda Aceh menggunakan Algoritma C4.5
   2. Mampu menganalisis konsep kerja Algoritma C4.5 pada sistem informasi untuk memprediksi kelulusan mahasiswa.
   3. Merancang sebuah sistem informasi prediksi kelulusan mahasiswa menggunakan Algoritma C4.5 dan *framework Codeigniter.*

b. Manfaat Praktis
   1. Bagi Peneliti, dengan dilakukannya penelitian ini diharapkan dapat memperluas pengetahuan dan wawasan peneliti tentang penerapan Algoritma C4.5 dalam memprediksi kelulusan mahasiswa angkatan 2020.
   2. Bagi Program Studi, diharapkan dari penelitian ini dapat memberi informasi tentang hasil analisis prediksi kelulusan mahasiswa.
   3. Bagi Fakultas Sains dan Teknologi UIN Ar-Raniry Banda Aceh, diharapkan dapat memberikan informasi mengenai prediksi kelulusan mahasiswa angkatan 2020 yang dapat dijadikan acuan untuk memprediksi kelulusan mahasiswa angkatan berikutnya.

c. Manfaat Kebijakan
   1. Bagi Peneliti, dengan dilakukannya penelitian ini diharapkan dapat dijadikan sebagai bahan referensi untuk melakukan penelitian dengan topik terbaru atau mengembangkan topik sebelumnya sehingga dapat memberikan manfaat bagi para pembaca
   2. Bagi Program Studi, diharapkan dari penelitian ini dapat membuat sebuah kebijakan untuk dapat menyelesaikan permasalahan terkait keterlambatan kelulusan mahasiswa agar dapat menciptakan proses pembelajaran yang lebih baik.
   3. Bagi Fakultas Sains dan Teknologi UIN Ar-Raniry Banda Aceh, diharapkan dapat membuat sebuah kebijakan untuk membuat sistem

pendidikan dan pembelajaran yang aplikatif, kondisional, dan tepat sasaran untuk menghasilkan lulusan mahasiswa yang berkualitas.

## I.6 Sistematika Penulisan

Dalam penyusunan tugas akhir, sistematika penulisan terbagi menjadi 5 bab. Berikut ini merupakan gambaran umum mengenai pokok pembahasan yang akan dibahas pada tiap-tiap bab :

**BAB 1 PENDAHULUAN**

Pada bab ini akan diuraikan latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

**BAB II KAJIAN PUSTAKA**

Pada bab ini akan diuraikan secara singkat mengenai penelitian terdahulu, dan teori-teori yang berkaitan dengan judul.

**BAB III METODE PENELITIAN**

Pada bab ini menjelaskan tentang metode-metode yang digunakan dalam penelitian meliputi metode pengumpulan data, pengolahan data, analisis, dan pengujian.

**BAB IV HASIL DAN PEMBAHASAN**

Pada bab ini akan diuraikan hasil dan pembahasan dari permasalahan mengenai prediksi kelulusan mahasiswa Fakultas Sains dan Teknologi UIN Ar-Raniry Banda Aceh.

**BAB V KESIMPULAN & SARAN**

Pada bab ini menjelaskan kesimpulan dari pengklasifikasian kelulusan mahasiswa dengan Algoritma *C4.5* serta saran bagi penelitian selanjutnya.

# BAB II

# KAJIAN PUSTAKA

## II.1 Penelitian Terdahulu

Penelitian yang terkait pada penelitian ini dengan menggunakan metode data mining, memerlukan beberapa referensi dan perbandingan literatur yang berkaitan dengan tema penulisan untuk terhindar dari duplikasi dan mengembangkan suatu hal yang berbeda pada peelitian ini, seperti penelitian yang dilakukan oleh F.F.Anwar, A.I. Jaya, dan M. Abu dengan 16 atribut dan 1 atribut target yang diuji menggunakan 75 data *testing*. Diperoleh hasil 69 data mahasiswa diprediksi benar, dan 6 data mahasiswa diprediksi salah dengan hasil nilai *accuracy* 92%, *error rate* 8%, nilai *AUC* sebesar 0,7138 (Anwar, F.F dkk, 2022). Penelitian lainnya yang dilakukan Susi Mashlahah yang berjudul " Prediksi Kelulusan Mahasiswa Menggunakan Metode *Decision Tree* Dengan Penerapan Algoritma C4.5" memperoleh hasil akurasi kecocokan pada sistem mencapai 82,79% dengan atribut asal daerah, jenis sekolah, jalur masuk, pengalaman pesantren, IPK, dan IP semester 1 sampai 5 (Mashlahah Susi, 2013).

Prediksi kelulusan lain yang menggunakan algoritma C4.5 diteliti oleh Lydia Yohana Lumban Gaol, M.Sofii, dan Dedi Suhendro terhadap mahasiswa Stikom Tunas Bangsa Prodi Sistem Informasi memperoleh *accuracy* 90.00%, *precision* 91.38% dan *recall* 98.15%. Hal ini menunjukkan atribut IPK memiliki pengaruh paling tinggi dalam menentukan kelulusan mahasiswa (Gaol, Lydia Yohana Lumban dkk, 2021). Implementasi penelitian lainnya yang dilakukan oleh Abdul Rohman dan Anief Rufiyanto mengenai Prediksi Kelulusan Mahasiswa Menggunakan *Decision Tree* Di Universitas Pandanaran dengan menggunakan pengujian *K-fold Cross Validation*, validasi hasil dengan *confusion matrix*, dan kurva *ROC* dengan *AUC*. Dari penelitian tersebut, diperoleh hasil nilai akurasi 65,98% dan nilai *AUC* 0,874 dengan klasifikasi data baik (Rohman, Abdul & Rufiyanto, Anief, 2019).

Adapun penelitian prediksi kelulusan yang dilakukan oleh Lastri Yuningsih, Iwan Rizal Setiawan, dan Asril Adi Sunarto yang berjudul "Rancangan Aplikasi

Prediksi Kelulusan Siswa Menggunakan Algoritma C4.5 " memperoleh hasil dari uji coba menggunakan 104 data sampel dari siswa lulusan SMK Al-Mizab tahun 2018/2019, indeks prestasi *gain* tertinggi berada di semester 2 dengan nilai *gain* 0,31667 (Yuningsih, Lastri dkk, 2020) . Selain itu, penelitian lainnya yang dilakukan Ade Fatma Ayu Rahman, Sorikhi, da, Sri Wartulax mengenai prediksi kelulusan mahasiswa di Universitas Peradaban menggunakan algoritma C4.5 memperoleh hasil analisis yang menunjukkan bahwa algoritma C4.5 dapat memprediksi dengan nilai *accruracy* sebesar 88,74%, *precision* sebesar 91,79%, dan *recall* sebesar 95,34% (Rahman, Ade Fatma Ayu dkk, 2020).

Penerapan algoritma C4.5 lainnya untuk memprediksi kelulusan juga dilakukan oleh Ratna Puspita Sari Putri, dan Indra Waspada dapat disimpulkan bahwa atribut yang paling dominan dalam kelulusan mahasiswa adalah IPK, *TOEFL*, asal daerah, dan jenis kelamin. Penggunaan nilai *confidence* 0,25 meningkatkan akurasi lebih baik daripada nilai *Confidence* 0,4 (Putri, Ratna Puspita Sari & Waspada, Indra, 2018). Implementasi tersebut juga sama seperti yang dilakukan Sarah Novia Hermawanti, Asriyanik, dan Asril Adi Sunarto memiliki *accuracy* kecocokan sebesar 68,42%, *precision* 42,86%, dan *recall* 60%. Sistem yang dibuat menggunakan algoritma C4.5 berhasil memprediksi kelulusan (Hermawanti, Sarah Novia & Sunarto, Asril Adi, 2019).

Implementasi algoritma C4.5 yang dirancang, dan dikembangkan untuk memprediksi penentuan penerima beasiswa dilakukan oleh Abdurraghib Segaf Suweleh, Dyah Susilowati, dan Hairani dengan judul penelitian "Aplikasi Penentuan Beasiswa Menggunakan Algoritma C4.5 ". Aplikasi tersebut dibangun menggunakan metode *Waterfall*, *PHP*, *MySQL*, dan *Black Box* sebagai pengujian sistem memperoleh klasifikasi penerima beasiswa dengan tingkat akurasi mencapai 92%, spesifitas 92.3%, dan sensifitas 91.6% serta hasil pengujian sistem yang valid (Suweleh, Abdurraghib Segaf dkk, 2020).

Selain itu, algoritma C4.5 juga dapat diterapkan pada aplikasi untuk memprediksi kelulusan mahasiswa Fakultas Ilmu Komputer Universitas Brawijaya yang dilakukan oleh Febrian Pandu Widhianto, Ahmad Afif Supianto, dan Nanang Yudi Setiawan. Aplikasi tersebut dibangun menggunakan *HTML, CSS, PHP,*

*Javascript* dengan *framework Laravel*. Atribut yang berpengaruh dalam memprediksi kelulusan yaitu IPK lulus semester 4 dan IP beban semester 2. Dari penelitian tersebut memperoleh nilai akurasi menggunakan algoritma C4.5 sebesar 90.9% dan *output* yang dihasilkan berupa tampilan beberapa grafik pada *dashboard* dengan memuat presentase kelulusan, rekap kelulusan, dan detail nilai (Widhianto, Febrian Pandu dkk, 2019).

## II.2 Kajian Teoritis

Dibawah ini merupakan kajian terkait teori-teori yang digunakan pada penelitian tugas akhir :

### II.2.1 Data Mining

Menurut (Santoso, 2007), Data mining dapat disebut juga sebagai *Knowledge Discovery in Databases* (KDD). KDD merupakan proyek menggabungkan pengumpulan, analisis data, dan riset historis untuk mengungkap pola, koneksi atau hubungan lain diantara kumpulan data yang besar. Data mining adalah serangkaian proses yang dilakukan untuk mendapatkan pengetahuan atau pola dari kumpulan data yang memiliki beberapa metode seperti deskripsi, estimasi, prediksi, klasifikasi, dan pengklusteran.

### II.2.2 Fungsi Data Mining

Data mining memiliki tujuh fungsi dasar sebagai berikut :

1. ***Prediction* (Prediksi)** : Menemukan pola yang hilang dari sebuah data menggunakan analisis regresi untuk menemukan data yang tidak tersedia.
2. ***Sequencing* (Sekuensi)** : Bentuk jamak dari asosiasi yang berfungsi untuk memastikan hubungan berbeda dari kumpulan data yang didapat.
3. ***Classification* (Klasifikasi)** : Proses penemuan model yang berfungsi dalam menyimpulkan beberapa definisi karakteristik pada suatu grup atau kelompok.
4. ***Association* (Asosiasi)** : Identifikasi hubungan dari kejadian-kejadian yang telah terjadi dengan menyimpulkan definisi karakteristik sebuah data.

5. ***Clustering* (Pengelompokkan)** : Proses pengelompokkan sejumlah data atau objek yang memiliki kemiripan atribut.
6. ***Forecasting*** : Sistem perkiraan nilai prediksi menurut pola-pola pada gabungan data.
7. ***Description* (Deskripsi)** : Metode yang dapat membantu mendapatkan bentuk spesifik yang tersembunyi pada sebuah data.

II.2.3 Ruang Lingkup Data Mining

Data mining memiliki empat ruang lingkup dalam keilmuan, diantaranya:

1. ***Predictive Modelling*** : Ilmu yang menrangkum perancangan dan pengembangan algoritma untuk memungkinkan komputer mengembangkan perilaku berlandaskan data empiris dan penambangan data.
2. ***Cluster Analysis*** : Proses untuk mengetahui kelompok objek yang cocok satu sama lain tetapi beda dengan objek di kelompok lainnya.
3. ***Association Analysis*** : Untuk menemukan hubungan yang tidak terlihat dalam kumpulan data yang berukuran besar.
4. ***Anomaly Detection*** : Pengidentifikasian data yang tidak umum berupa *outlier*, perubahan atau deviasi.

II.2.4 Proses Data Mining

Tahapan proses data mining sebagai berikut :

1. ***Data Cleansing*** : Proses pembersihan data yang berubah-ubah dan *noise*.
2. ***Data Integration*** : Proses menggabungkan data dari sumber yang berbeda.
3. ***Data Selection*** : Proses penyeleksian data yang relevan untuk analisis dan didapatkan dari koleksi data yang ada.
4. ***Data Transformation*** : Proses transformasi data yang terpilih ke dalam bentuk *mining procedure*.
5. ***Data Mining*** : Proses untuk mengekstrak berbagai pola potensial untuk mendapatkan sejumlah data yang berguna.

6. ***Pattern Evolution*** : Pengolahan informasi yang sebelumnya ditemukan dengan cara mengidentifikasi berdasarkan *measure* yang diberikan.
7. ***Knowledge Presentation*** **:** Membantu pengguna dalam menginterpretasikan hasil dari pencarian data.

II.2 .5 Penerapan Data Mining

Penerapan data mining di kehidupan sehari-hari sebagai berikut :

1. Telekomumikasi

   Penerapan pengumpulan data pada perusahaan telekomunikasi dapat diterapkan untuk melihat jutaan transaksi yang masuk.
2. Keuangan

   Penerapan data mining dalam sektor keuangan dapat digunakan oleh berbagai bank untuk mengumpulkan informasi dan mudah mendeteksi transaksi keuangan yang mencurigakan.
3. Pendidikan

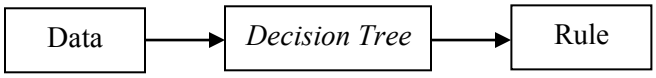
   Pada bidang pendidikan, data mining diterapkan untuk menemukan informasi yang berguna untuk kemajuan dalam bidang pendidikan seperti melakukan prediksi terhadap hasil ujian nasional siswa.
4. Pertanian

   Pada sektor pertanian, data mining dapat digunakan untuk mewujudkan ketahanan pangan tingkat nasional dengan cara pengelompokkan provinsi di Indonesia berdasarkan hasil panen.

II.3 *Decision Tree* (Pohon Keputusan)

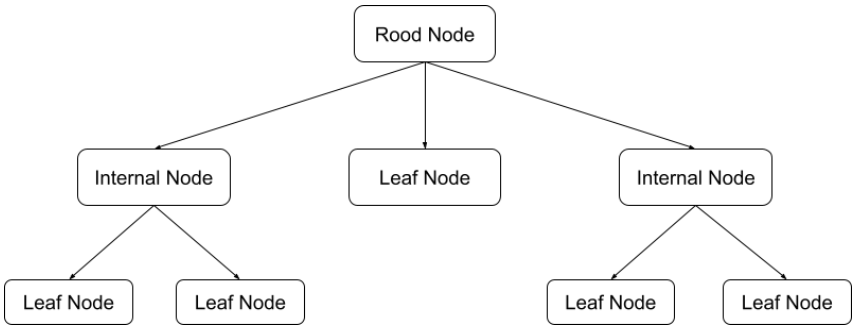
*Decision Tree* atau Pohon keputusan adalah metode klasifikasi dan prediksi untuk mengubah fakta menjadi pohon keputusan yang merepresentasikan aturan, mengeksplorasi data, dan menemukan hubungan tersembunyi antara sejumlah calon variabel input dengan variabel target. Aturan tersebut nantinya dapat ditampilkan dalam bentuk bahasa basis data, salah satunya bahasa *Structured Query Languange* (*SQL*) untuk menemukan *record* pada kategori tertentu. *Decision Tree* dapat disebut juga sebagai diagram alir yang memiliki bentuk seperti struktur pohon dimana setiap internal node menyatakan pengujian terhadap suatu atribut dan setiap cabang menyatakan output dari pengujian tersebut serta node daun (*leaf* node) (Mashlahah Susi, 2013).

Data pada pohon keputusan biasa dinyatakan dalam bentuk tabel dengan atribut dan *record*. Atribut menyatakan suatu parameter sebagai kriteria dalam pembentukan pohon keputusan dan memiliki nilai-nilai yang disebut dengan *instance.* Konsep yang dimiliki oleh pohon keputusan bertujuan untuk mengubah data menjadi pohon dan aturan-aturan keputusan. Proses pada pohon keputusan untuk mengubah bentuk data (tabel) menjadi model pohon, mengubah model pohon menjadi *rule*, dan menyederhanakan *rule*. Langkah perhitungan yang dilakukan yaitu menghitung nilai *entropy* total dari jumlah data yang dijadikan sampel, mengelompokkan variabel, dan menghitung nilai *gain* pada setiap atribut. Tahapan perhitungan tersebut akan berulang secara *continue* hingga beberapa tingkatan sampai mencapai nilai akhir yaitu keputusan "Yes" dan "No", tetapi pada penelitian ini keputusan yang digunakan "Lulus lebih awal", "Lulus tepat waktu", dan "Lulus terlambat".




**Gambar II. 1** Konsep Pohon Keputusan (Sumber : Data di Olah, 2023)

Pada Gambar II.1 menjelaskan tentang konsep dari pohon keputusan. Data yang telah diperoleh akan dibentuk menjadi pohon keputusan dengan menghasilkan beberapa *rule* atau aturan-aturan.




**Gambar II. 2** Model Keputusan *Decision Tree* (Sumber : Data di Olah, 2023)

Berdasarkan Gambar II.2 diatas, diperoleh sebuah model keputusan. Atribut dengan nilai *gain* tertinggi akan menjadi *rood* node. Setelah nilai *gain* pertama di

dapat, akan dilakukan perhitungan ulang untuk mencari nilai *gain* tertinggi dari atribut lainnya. Sehingga, terbentuk sebuah pohon keputusan yang memiliki nilai.

Pohon keputusan memiliki manfaat untuk mem-*breakdown* proses pengambilan keputusan yang kompleks menjadi lebih sederhana sehingga dapat lebih menginterpretasikan solusi dari permasalahan. Ada beberapa keunggulan dari metode *decision tree* yaitu :

a. Memiliki kecepatan yang relatif lebih cepat
b. Dapat diubah menjadi rule klasifikasi dengan mudah dan sederhana
c. Dapat menggunakan *query SQL* untuk mengakses *database*
d. Dapat membandingkan tingkat akurasi dengan metode lainnya

## II.4 Algoritma C4.5

Algoritma C4.5 merupakan teknik yang digunakan untuk melakukan kategorisasi bersifat prediktif dan termasuk algoritma pohon keputusan yang menggunakan rancangan perolehan informasi untuk menentukan pembagian yang tertinggi.

Terdapat beberapa perbedaan antara prinsip dasar kerja algoritma C4.5 dan algoritma ID3, yaitu:

1. Menangani atribut yang berkelanjutan.
2. Menangani atribut yang kosong (*Missing Value*).
3. Pemangkasan pohon keputusan.
4. Pemilihan atribut menggunakan *Gain Ratio.*

## II.5 *Classification* (Klasifikasi)

Klasifikasi adalah bagian dari suatu data mining dengan berbagai macam model diterapkan ke dalam banyak kasus dengan menghasilkan tingkat akurasi yang berbeda-beda. Algoritma yang sering digunakan adalah *C4.5, Logistic Regression (LR), K-NN, Naïve Bayes (NB), Neural Network (NN),* dan *Support Vector Machine (SVM).* Algoritma C4.5 merupakan sebuah metode klasifikasi untuk memprediksi sebuah label atau output yang sebelumnya telah diinput ke dalam data *training.* Klasifikasi data terdiri dari dua proses. Pertama, *learning* (fase *training*) yaitu algoritma yang dibuat untuk menganalisa data *training* kemudian

direpresentasikan ke dalam bentuk *rule* klasifikasi. Kedua, klasifikasi dimana data *testing* digunakan untuk memperkirakan akurasi dari *rule* klasifikasi (Rohman, Abdul. 2019)

Proses klasifikasi didasarkan pada empat komponen, diantaranya sebagai berikut :

a. ***Class***. Variabel dependen berupa kategorikal yang merepresentasikan label pada objek.

   Contoh : resiko penyakit jantung, resiko kredit, *customer loyalty*, jenis gempa.

b. ***Predictor***. Variabel *independent* yang direpresentasikan oleh karakteristik (atribut) data.

   Contoh : merokok, minum alcohol, tekanan darah, tabungan, asset, gaji.

c. ***Training dataset***. Sebuah set daya yang memiliki nilai dari komponen *class* dan *predictor* untuk menentukan kelas yang cocok menurut *predictor.*

d. ***Testing dataset***. Data baru yang akan diklasifikasikan oleh model yang telah dibuat dan akurasi klasifikasi dievaluasi.

### *II.6 Confusion Matrix*

*Confusion matrix* atau disebut juga dengan *error matrix* merupakan alat yang terbuat dari elemen matriks 2x2 dan digunakan untuk mendapatkan jumlah kumpulan data yang memenuhi kriteria untuk klasifikasi aktif dan tidak aktif pada masing-masing dua algoritma yang digunakan (Mustika, Yunita Ardilla. 2021).

Penggunaan *confusion matrix* dapat memberikan manfaat antara lain :

1) Menunjukkan cara kerja model saat membuat prediksi.
2) *Confusion matrix* dapat memberikan informasi rincian jenis kesalahan yang dibuat model.
3) Masing-masing kolom dari tabel matiks konfusi menampilkan contoh kelas prediksi.
4) Masing-masing baris dari tabel matriks konfusi menggantikan contoh kelas aktual.

Terdapat empat istilah sebagai hasil dari proses pendeteksian objek, yaitu :

- *True Positive* (TP), diartikan memprediksi positif dan benar
- *True Negative* (TN), diartikan memprediksi negatif dan benar
- *False Negative* (FP), diartikan memprediksi positif dan salah
- *False Negative* (FN), diartikan memprediksi negatif dan salah.

Terdapat cara yang lebih mudah untuk istilah-istilah pada *confusion matrix*, yaitu :

a. Jika awalan bernilai *True,* maka prediksi benar, meskipun diprediksi terjadi atau tidak terjadi.
b. Jika awalan bernilai *False*, maka prediksi salah.
c. Nilai positif dan negatif merupakan hasil perkiraan dari model.

Terdapat tiga *performance metric* populer yang telah umum digunakan, yaitu akurasi, *presisi, recall, dan f-measure* :

**1) *Confusion Matrix* untuk Menghitung *Accuracy***

Akurasi adalah kedekatan nilai yang di estimasi dengan nilai sebenarnya. Keakuratan dari sebuah bentuk klasifikasi menggambarkan keakuratan model yang dibangun dan dapat mengklasifikasikan dengan benar. Adapun rumus persamaan akurasi dapat dilihat pada persamaan 2.1.

Rumus :

$$Accuracy = \frac{Jumlah\ objek\ terdeteksi\ benar}{Jumlah\ keseluruhan\ objek\ yang\ terdeteksi} \text{ x } 100\% \text{..... Persamaan (2.1)}$$

**2) *Confusion Matrix* untuk Menghitung *Precision***

Presisi adalah rasio prediksi positif dibandingkan dengan semua hasil yang diprediksi positif. Nilai presisi menggambarkan tingkat kesesuaian antara data yang diprediksi dan hasil aktual. Presisi mengukur berdasarkan semua kelas positif yang telah ditentukan kebenarannya. Adapun rumus persamaan presisi dapat dilihat pada persamaan 2.2.

Rumus :

$$Precision = \frac{TP}{(TP + FP)} \text{ x } 100\% \text{..... Persamaan (2.2)}$$

3) ***Confusion Matrix* untuk Menghitung *Recall***

*Recall* adalah rasio prediksi benar positif dibandingkan dengan semua hasil yang diprediksi benar positif dan salah negatif. *Recall* menggambarkan keberhasilan model dalam menemukan kembali sebuah informasi. Selain itu, *recall* juga dikenal sebagai sensitivitas dalam klasifikasi *biner diagnostic*. Adapun rumus persamaan *recall* dapat dilihat pada persamaan 2.4.

Rumus :

$$Recall = \frac{TP}{(TP + FN)} \text{ x } 100\% \text{ .....Persamaan (2.3)}$$

4) ***Confusion Matrix* untuk Menghitung F- *Measure***

F- *score* atau disebut juga F- *measure* adalah ukuran akurasi suatu pengujian dimana dihitung dari hasil pengujian presisi dan *recall*. Rentang dari nilai f-*measure* biasanya antara 0 sampai dengan 1 dan tidak dijadikan bentuk persentase (meskipun beberapa penelitian menulisnya dalam bentuk persentase). Semakin mendekati nilai 1, maka semakin mungkin untuk menemukan kembali data dengan atribut yang serupa.

## *II.7 Website*

*Website* atau dapat disebut juga dengan *WWW* (*World Wide Website*) adalah sebuah media informasi yang dapat diakses melalui internet menggunakan metode penentuan alamat yang unik. *Website* berisi perpaduan antara teks, suara, hypermedia, dan grafis. *Website* dapat mempermudah penggunanya untuk melakukan hubungan jarak jauh dengan antarmuka grafis yang dimiliki (Riyadi, Firman Azhar, 2020).

## II.8 Basis Data

Basis data atau yang sering disebut dengan *database* merupakan kumpulan tabel yang berisi sekumpulan data sebagai sumber informasi yang disimpan dalam media penyimpanan secara digital untuk memudahkan aktivitas memperoleh informasi yang dibutuhkan (TJ Sitinjak, Daniel Dido Jantce dkk, 2020).

## *II.9 MySQL*

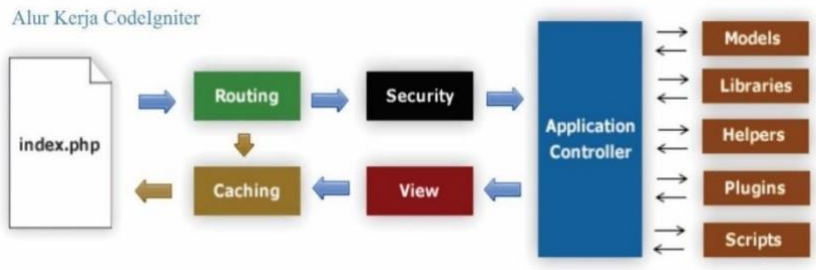
*MySQL* (*My Structure Query Languange*) merupakan aplikasi *DBMS* atau *software database* yang digunakan untuk mengolah basis data. *MySQL* bersifat *open source* dan menggunakan bahasa *SQL* (*Structure Query Languange*). *MySQL* berfungsi untuk menyimpan data dalam bentuk tabel yang saling berhubungan dan sering di *bundling* dengan *web server* sehingga proses instalasinya menjadi lebih mudah. Selain itu, *MySQL* juga dapat dijalankan diberbagai platform seperti *Windows, Linux,* dan lain sebagainya ( Maman dkk, 2020)

## *II.10 PhpMyadmin*

*PhpMyadmin* merupakan sebuah aplikasi berbentuk halaman situs pada *web server* yang berfungsi untuk memudahkan manajemen *MySQL*. *PhpMyadmin* bersifat *open source* yang dapat membuat *database*, tabel, *insert*, *delete,* dan *update* data dengan *GUI* tanpa perlu menulis perintah *SQL* secara manual (Suwita, Jaka, 2020)

## *II.11 Codeigniter*

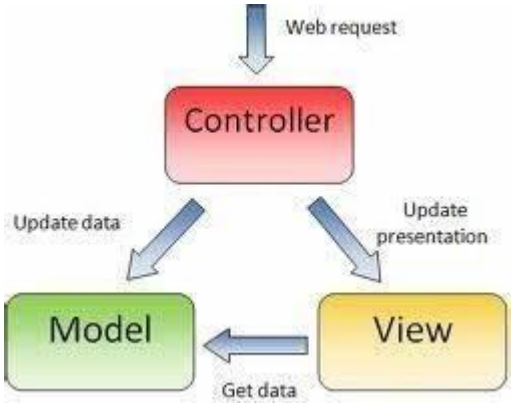
*Codeigniter* atau yang sering disebut dengan CI adalah *framework* pengembangan aplikasi (*Application Development Framework*) yang bersifat *open source* menggunakan *PHP* dengan model *MVC* (*Model, View, Controller*) untuk membangun situs web yang dinamis. *Codeignitier* pertama kali dibuat oleh Rick Ellis pada tanggal 28 Februari 2006. Tujuannya untuk memudahkan pengembang web membuat aplikasi dengan cepat dibandingkan dengan membuatnya dari awal (Daniel Dido dkk, 2020)




**Gambar II. 3** Alur Kerja *Codeigniter* (Sumber : www.1.bp.blogspot.com/)

Berdasarkan gambar alur kerja *codeignitier* II.3, berikut ini merupakan fungsi dari alur-alur tersebut :

1. File *index.php* berfungsi sebagai *front controller.*
2. *Router*, berfungsi memeriksa *HTTP request* untuk dapat menentukan apa yang harus dilakukan.
3. Jika file *cache* ada, akan dikirim langsung ke *browser* melewati eksekusi sistem normal.
4. *Security*, sebelum *controller* aplikasi dimuat, *HTTP request* dan setiap data pengguna disubmit akan diseleksi terlebih dahulu untuk keamanannya.
5. *Controller* berfungsi untuk memuat model, *library* utama, *helper,* dan setiap *resource* yang diperlukan untuk memproses permintaan khusus.
6. *View*, proses *render* yang dikirim ke *web browser* dengan tujuan untuk dapat dilihat. Jika *caching* diaktifkan, *view* akan *dicache* terlebih dahulu sehingga dapat dilayani pada permintaan berikutnya.




**Gambar II. 4** Konsep MVC *Codeigniter* (Sumber : www.maniacms.web.id/)

Berikut ini merupakan fungsi dari *MVC* (*Model, View, Controller*) pada Gambar II.4 :

1. **Model**, digunakan sebagai presentasi *database* yang menyimpan perintah-perintah *Query SQL* seperti *insert*, edit, *delete*, dan *select*.

2. ***View*** merupakan halaman khusus yang digunakan untuk menampilkan informasi kepada *client* dengan segala macam permintaan yang dikelola oleh *controller* dan model akan dikembalikan kepada *view* sesuai hasil permintaan yang diinginkan.
3. ***Controller***, digunakan sebagai pengendali (*control*) antara *view* dan model melalui permintaan dari *HTTP*.

### *II.12 PHP*

*PHP* merupakan singkatan dari *Hypertext Preprocessor* yaitu bahasa pemrograman web *server side scripting* yang bersifat *open source*. *PHP* digunakan untuk membuat halaman web yang dinamis. Bahasa pemrograman *PHP* bisa dapat dikombinasikan dengan bahasa lain, seperti *HTML, CSS, Javascript*, atau gabungan keempatnya. *PHP* mempunyai ekstensi sendiri yakni ekstensi .php. kelebihan dari bahasa *PHP* adalah mudah untuk dipelajari, dan memiliki pengembangan serta pembaruan lebih cepat dibandingkan bahasa pemrograman lainya. Akan tetapi, bahasa *PHP* juga memiliki kekurangan salah satunya tidak memiliki tipe data sehingga sering menimbulkan *bug* (Daniel Dido, dkk, 2020).

### II.13 *UML* (*Unifield Modelling Languange*)

*UML* merupakan singkatan dari *Unifield Modelling Languange* yaitu sebuah metode pemodelan yang digunakan untuk memvisualisasikan perancangan sistem berorientasi objek. *UML* terdiri dari *Structural Classification, Dynamic Behavior,* dan model management. *UML* menginterpretasikan diagram sebagai *Use case diagram, Class diagram, Activity diagram, Sequence fiagram, Statechart diagram, Collaboration diagram, Component diagram, dan Deployment diagram.* (Pardosi, Wilson dkk, 2020).

### *II.14 Framework*

*Framework* merupakan sebuah kerangka kerja yang dapat memudahkan pengembang aplikasi untuk membangun *software* sesuai dengan bahasa pemrograman. Untuk bahasa pemrograman *PHP* biasanya terdiri dari sekumpulan folder yang berisi berbagai file *PHP* dimana file tersebut merupakan *Class Library, Helper, Plugin,* atau konfigurasi lainnya. *Codeignitier* adalah salah satu *framework* untuk membangun website yang sering digunakan menggunakan bahasa

pemrograman *PHP*. Tujuannya untuk memungkinkan pengembangan proyek lebih cepat dari pada penulisan kode terstruktur. Dengan antarmuka yang sederhana, struktur logika yang digunakan untuk mengakses *library* membuat *Codeignitier* mudah dipelajari (Damanik, Burhanuddin dkk, 2020).

### *II.15 Bootstrap*

*Bootstrap* adalah *Framework HTML, CSS*, dan *JavaScript* yang bersifat *open source* dan memiliki fungsi untuk mendesain *website responsive* dengan mudah. *Bootstrap* diciptakan oleh Mark Otto dan Jacob Thornton dari Twitter pada tahun 2011 dengan nama *Twitter Blueprint*. Menurut Alatas (2013:2) dalam bukunya yang berjudul *Responsive Web Design* dengan *PHP* dan *Bootstrap*, mengemukakan bahwa "*Bootstrap* merupakan *Framework* ataupun *tools* untuk membuat aplikasi *web* ataupun situs *web responsive* secara cepat, mudah dan gratis" (Daniel, Dido dkk, 2020).

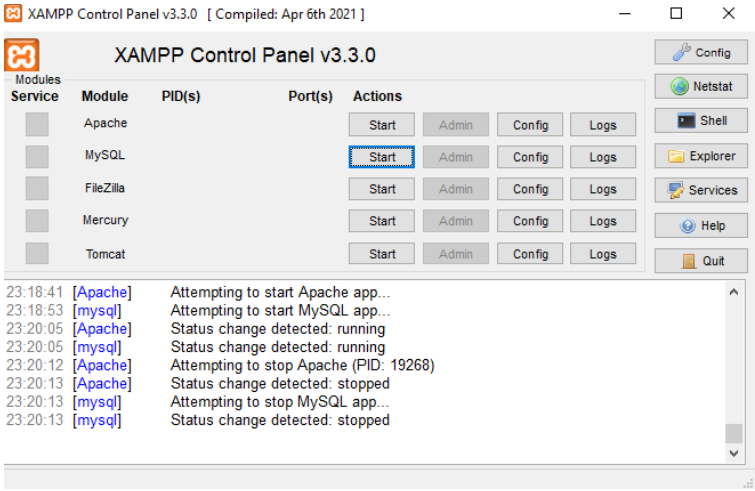
*Bootstrap* memiliki kegunaan dalam melakukan pengembangan *website* sebagai berikut :

1. Menciptakan *website Mobile Friendly* dengan sistem grid yang tidak membutuhkan waktu yg lama.
2. Memudahkan *resize* gambar hanya dengan menambahkan *class.img-responsive* ke dalam gambar.
3. Menambahkan elemen *website* tanpa ribet dengan berbagai elemen yang telah disediakan oleh *bootstrap*, seperti navigasi, menu *dropdown*, *thumbnail*, dan lain sebagainya.
4. Membuat *website* lebih interaktif dengan menggunakan *plugin custom* JQuery seperti popup, transisi, *image carousel*, dan lain-lain.

### *II.16 XAMPP*

*XAMPP* merupakan sebuah perangkat lunak yang berfungsi untuk menjalankan *website* berbasis *PHP* dan menggunakan *MySQL* untuk mengolah data di komputer. *XAMPP* memiliki peran penting sebagai *web server* pada komputer local. Selain itu, *XAMPP* dapat juga disebut sebagai *Cpanel Server Virtual* yang dapat membantu melakukan *preview* sehingga bisa dimodifikasi tanpa

menggunakan akses internet (TJ Sitinjak, Daniel Dido Jantce dkk, 2020). *XAMPP* yang digunakan pada penelitian ini adalah versi 3.3.0.



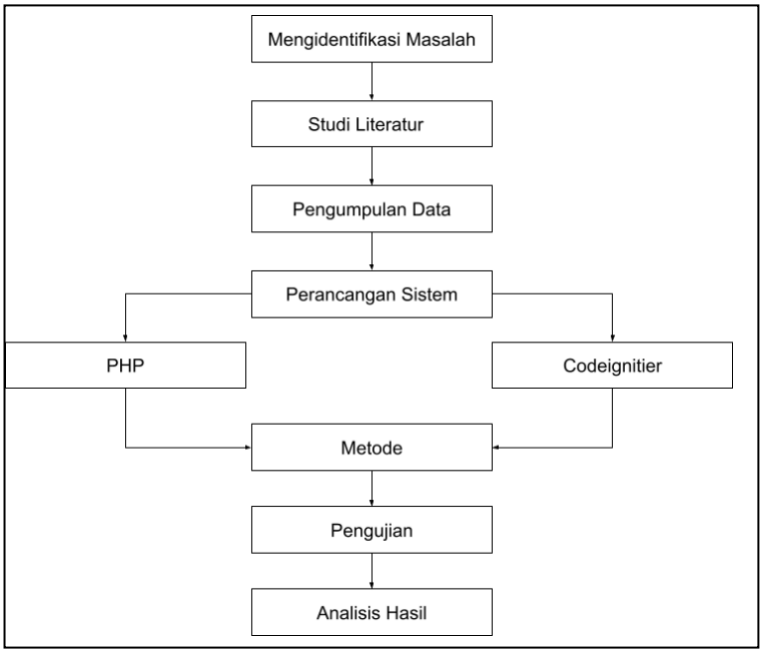

**Gambar II. 5** *XAMPP* (Sumber : Data di Olah, 2023)

Menurut Betha Sidik (2018:6), *XAMPP* merupakan singkatan huruf dari :

1. **X**: Program dapat dijalankan di berbagai sistem operasi, seperti *Windows, Linux, Mac OS,* dan Solaris.
2. **A** (***Apache***) : Aplikasi *web server* yang bertugas untuk menghasilkan halaman web kepada user berdasarkan kode *PHP* yang ditulislan oleh pembuat halaman web.
3. **M** (**MySQL**) : Aplikasi *database server* yang digunakan untuk membuat dan mengelola *database* serta memiliki fitur *insert, delete, update* yang berada dalam *database.*
4. **P** (**PHP**) : Bahasa pemrograman web yang digunakan untuk membuat web *server-side scripting* dan *web dinamis*. Selain itu, *PHP* juga dapat digunakan untuk mendukung pengelolaan sistem *database Oracle, Microsoft Access, Interbase, d-base, PostgreSQL,* dan lain-lain.
5. **P** (**Perl**) : bahasa pemrograman yang dapat digunakan untuk semua tujuan yang pertama kali dikembangkan oleh Larry Wall dan dirilis pada tanggal 18 Desember 1987. Perl tersedia diberbagai sistem operasi *UNIX, DOS, Windows, PowerPC, BeOS, VMS, EBCDIC,* dan *PocketPC.*

## II.17 Kerangka Penelitian

Secara garis besar, dalam merancang langkah-langkah pada penelitian ini adalah mengidentifikasi masalah, studi literatur, pengumpulan data, perancangan, metode, pengujian, dan analisa hasil. Metode algoritma *C4.5* yang digunakan bertujuan untuk melakukan kategorisasi bersifat prediktif yang menghasilkan sebuah keputusan berbentuk pohon keputusan. Pada Gambar II.6 dibawah ini memperlihatkan alur dari tahapan penelitian.



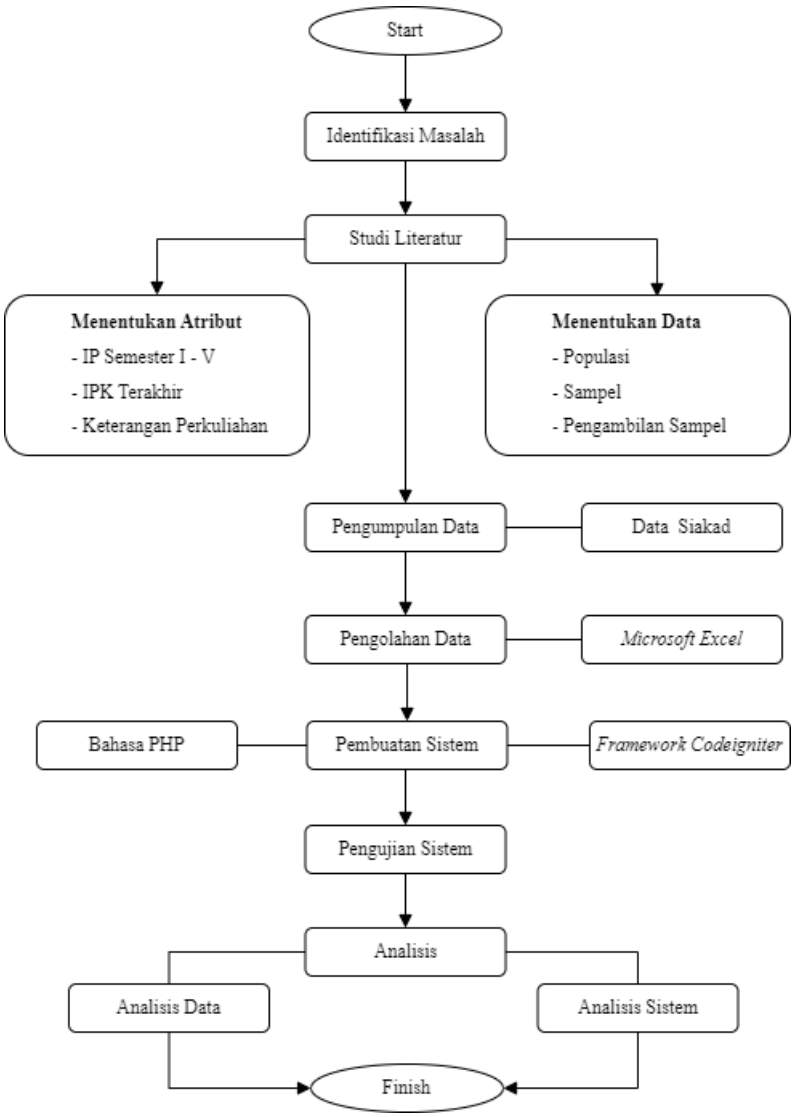

**Gambar II. 6** Kerangka Penelitian

# BAB III

# METODE PENGEMBANGAN PRODUK

## III.1 Metodologi Pengembangan Produk

Pada penelitian ini terdapat tujuh alur atau tahapan yang dilakukan untuk mengembangkan sebuah produk. Secara umum, tahapan penelitian ditampilkan pada Gambar III.1 :




**Gambar III. 1** Tahapan Penelitian

Adapun tahapan-tahapan yang dilakukan dalam penelitian ini adalah sebagai berikut :

1. **Identifikasi Masalah**

   Mengidentifikasi masalah merupakan tahapan penting pada saat melakukan penelitian dengan tujuan agar dapat mengetahui permasalahan.

2. **Studi Literatur**

   Studi literatur dilakukan untuk mencari referensi terkait penelitian terdahulu yang dapat menjadi penentu sumber data dan atribut yang digunakan dalam penelitian.

3. **Pengumpulan Data**

   Pengumpulan data dilakukan untuk memperoleh data yang dibutuhkan selama penelitian berlangsung. Data yang diperoleh merupakan data sekunder dari SIAKAD program studi Arsitektur, Kimia, Biologi, Teknik Lingkungan, dan Teknologi Informasi.

4. **Pengolahan Data**

   Pada tahap ini data akan dikelola menggunakan *Microsoft Excel* dengan membagi ke dalam 2 folder yaitu data mahasiswa angkatan 2019 dan data mahasiswa angkatan 2020.

5. **Pembuatan Sistem**

   Pembuatan sistem informasi prediksi kelulusan menggunakan bahasa pemrograman *PHP* dan *framework Codeignitier* serta *Sublime* sebagai *text editor*.

6. **Pengujian**

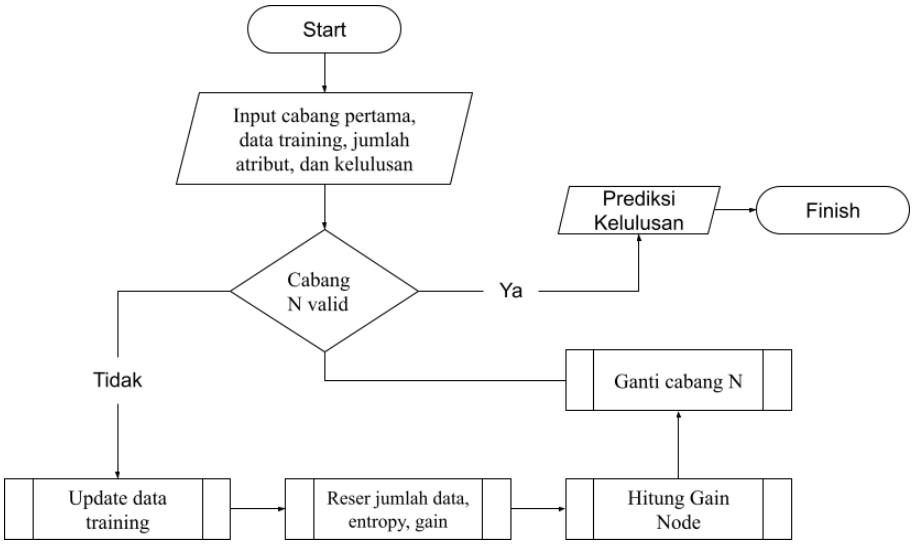
   Pengujian dilakukan untuk mengetahui keakuratan data, cara kerja, kelebihan, dan kekurangan sistem menggunakan *Confusion Matrix* dan *Black box testing.*

7. **Analisis**

   Analisis dilakukan untuk mengetahui tingkat keakuratan sistem dalma memberikan keputusan sesuai dengan permasalahan.

### III.1.1 Algoritma C4.5

Algoritma C4.5 merupakan struktur pohon yang memiliki simpul dan mendeskripsikan berbagai atribut, setiap cabang yang menggambarkan hasil dari atribut, dan setiap daun yang menggambarkan kelas.




**Gambar III.1 1** *Flowchart* Algoritma *C4.5*

Terdapat tujuh tahapan dalam membuat sebuah keputusan menggunakan algoritma C4.5*,* yaitu :

**Gambar III.1 2** Blok Diagram Algoritma *C4.5*

1. Menyiapkan dataset
2. Menghitung nilai *Entropy* dengan menggunakan rumus dari persamaan 3.1.

   Rumus :

   $$Entropy\ (S) = \sum_{i=1}^{n} -pi * log_2\ pi \text{ ..... Persamaan (3.1)}$$

   Keterangan :

   S = Himpunan kasus

   n = Jumlah partisi S

   $p_i$ = Proporsi $S_i$ terhadap S

3. Menghitung nilai *Gain* dengan menggunakan rumus dari persamaan 3.2.

   Rumus :

   $$Gain\ (S, A) = Entropy\ (S) - \sum_{i=1}^{n} \frac{|S_i|}{|S|} * Entropy\ (S_i) \text{.....Persamaan (3.2)}$$

   Keterangan :

   S = Himpunan kasus

   A = Fitur

   N = Jumlah partisi atribut A

   $|S_i|$ = Proporsi $S_i$ terhadap S

   $|S|$ = Jumlah kasus dalam S

4. Menghitung nilai *Split Info* dengan menggunakan rumus dari persamaan 3.3.

   Rumus :

   $$SplitInfo\ (S, A) = -\sum_{i=1}^{n} \frac{S_i}{S} log_2 \frac{S_i}{S} \text{ ..... Persamaan (3.3)}$$

   Keterangan :

   S = Ruang (data) *sample* yang digunakan untuk training

   A = Atribut

   $S_i$ = Jumlah *sample* untuk atribut I

5. Menjumlah nilai *Gain Ratio* dengan menggunakan rumus dari persamaan 3.4.

   Rumus :

$$GainRatio = \frac{Gain\ (A)}{SplitInfo\ (S,A)} \quad \text{Persamaan (3.4)}$$

6. Membuat cabang untuk tiap-tiap nilai
7. Ulangi proses untuk cabang hingga semua *node* partisi

## III.2 Waktu dan Tempat Penelitian

Penelitian **"**Penerapan Algoritma C4.5 Untuk Memprediksi Kelulusan Mahasiswa Fakultas Sains dan Teknologi (Studi Kasus: Teknologi Informasi, Teknik Lingkungan, Arsitektur, Kimia, Biologi)**"** dilaksanakan pada bulan Desember 2022 s/d Maret 2023. Adapun tempat pelaksanaan penelitian dilakukan di Fakultas Sains dan Teknologi UIN Ar-Raniry Banda Aceh.

## III.3 Perangkat Penelitian

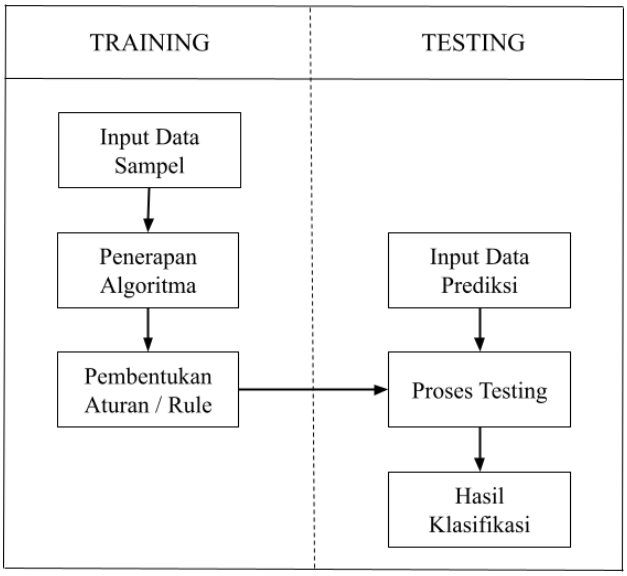
Dalam pembuatan sistem informasi prediksi kelulusan ini memerlukan perangkat keras, perangkat lunak, dan data sebagai penunjang agar sistem dapat berjalan sesuai dengan fungsinya.

**Tabel III. 1**. Perangkat penelitian

| No | Kebutuhan *Hardware* | No | Kebutuhan *Software* |
|---|---|---|---|
| 1 | Processor Core i3 | 1 | XAMPP Control Panel version 3.3.0 |
| 2 | Sistem Operasi Windows 10 (64-bit) | 2 | Apache version 2.4.53 |
| 3 | Free Hardisk minimal 20 GB | 3 | PHP version 7.4.29 |
| 4 | RAM 4 GB | 4 | *Framework Codeignitier* |
| 5 | Mouse | 5 | Sublime Text Editor 3 |
| 6 | Flashdisk | 6 | *Microsoft Excel* |
| | | 7 | *Draw.io.com* |
| | | 8 | *Wireframepro.mockflow.com* |

## III.4 Deskripsi Sistem

Sistem informasi yang akan dibangun menggunakan data training sebesar 80% dan data testing sebesar 20% (Gholamy,Afshin dkk, 2018). Secara garis besar memiliki alur seperti blok diagram dibawah ini:




**Gambar III. 2** Diagram Blok Sistem (Sumber : Data di Olah, 2023)

Berdasarkan diagram blok sistem pada Gambar III.2 diatas, sistem dimulai dengan dua tahapan, yaitu (Hermawanti et al., 2019) :

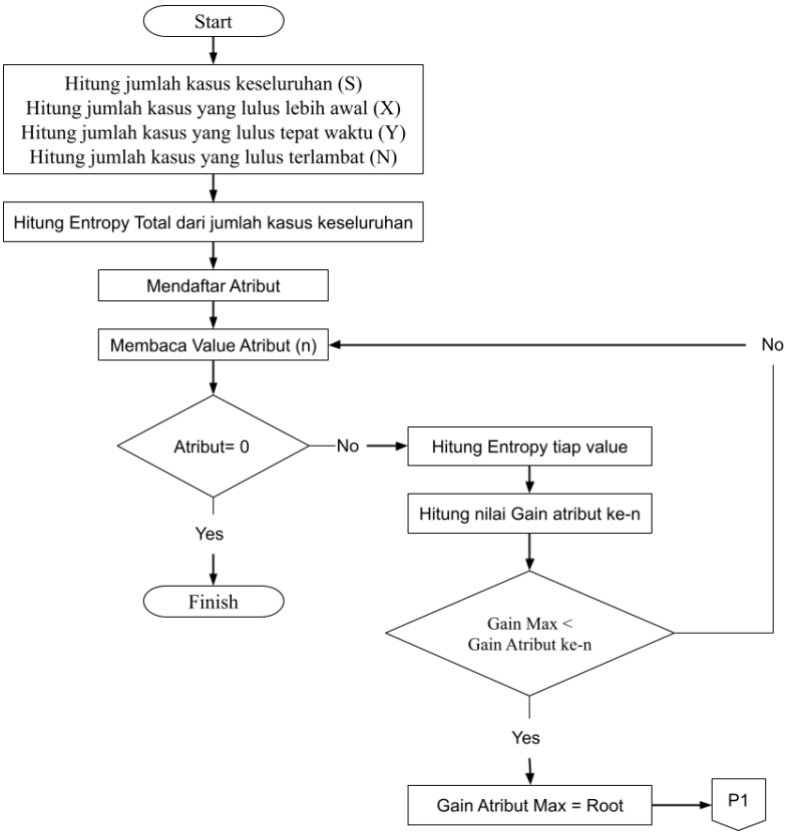
1. Proses *Training*

   Proses memasukkan data sampel ke dalam tabel yang dipersiapkan untuk melakukan proses perhitungan. Tabel tersebut meliputi atribut, jumlah data keseluruhan, dan jumlah data yang sudah terklasifikasi berdasarkan target yang telah ditentukan. Pada kasus ini ada tiga target yaitu lulus lebih awal, lulus tepat waktu, dan lulus terlambat, serta kolom nilai *Entropy* dan *Gain*. Tahap selanjutnya penerapan algoritma C4.5 untuk menghitung nilai *Entropy* dan *Gain* pada setiap atribut untuk dijadikan bentuk *Tree* yang akan diterapkan pada proses *testing*.

2. Proses *Testing*

   Proses memasukkan data uji atau data prediksi dengan menggunakan atribut yang sesuai dengan atribut pada proses training. Setiap data atribut akan dibandingkan dengan aturan yang telah terbentuk pada perhitungan data *training*. Setelah itu, data akan diklasifikasikan dengan kondisi atribut lulus lebih awal, lulus tepat waktu, dan lulus terlambat.

Dibawah ini merupakan *flowchart* pembentuk pohon keputusan (*Decision Tree*) pada sistem informasi prediksi kelulusan (Mashlahah, 2013):

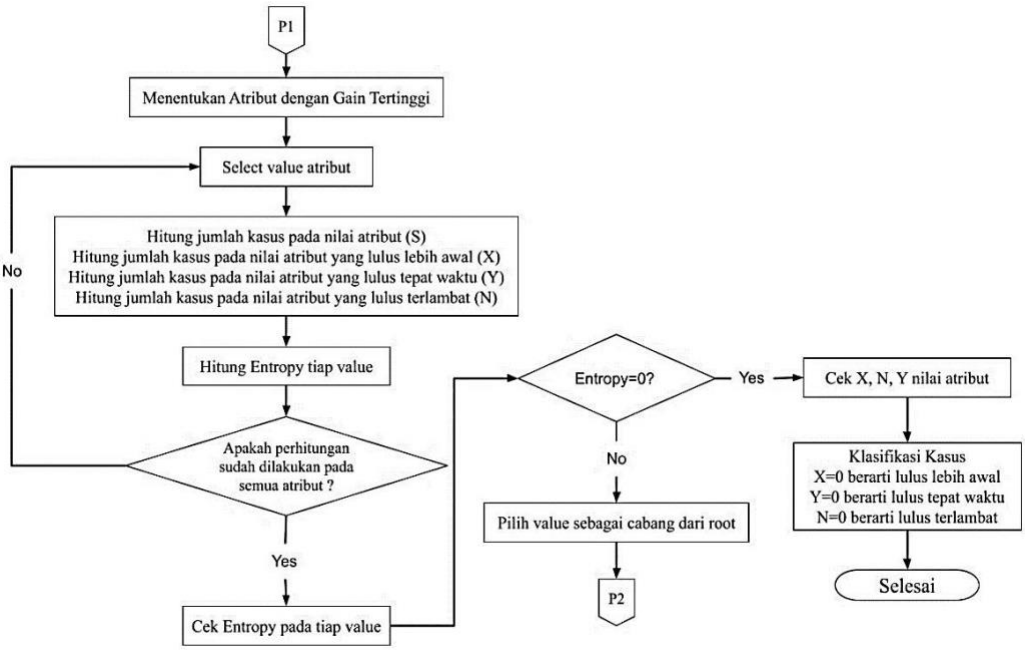



**Gambar III. 3** *Flowchart* Penentuan *Root* (Sumber : Data di Olah, 2023)

Berikut ini sistem yang berjalan pada saat penentuan *root* (akar) :

a. Sistem akan menghitung total nilai dari data *training*.
*b.* Setiap atribut akan dihitung nilai *Entropy* dan *Gain.*
c. Sistem akan membandingkan nilai *Gain* terbesar dari setiap atribut.

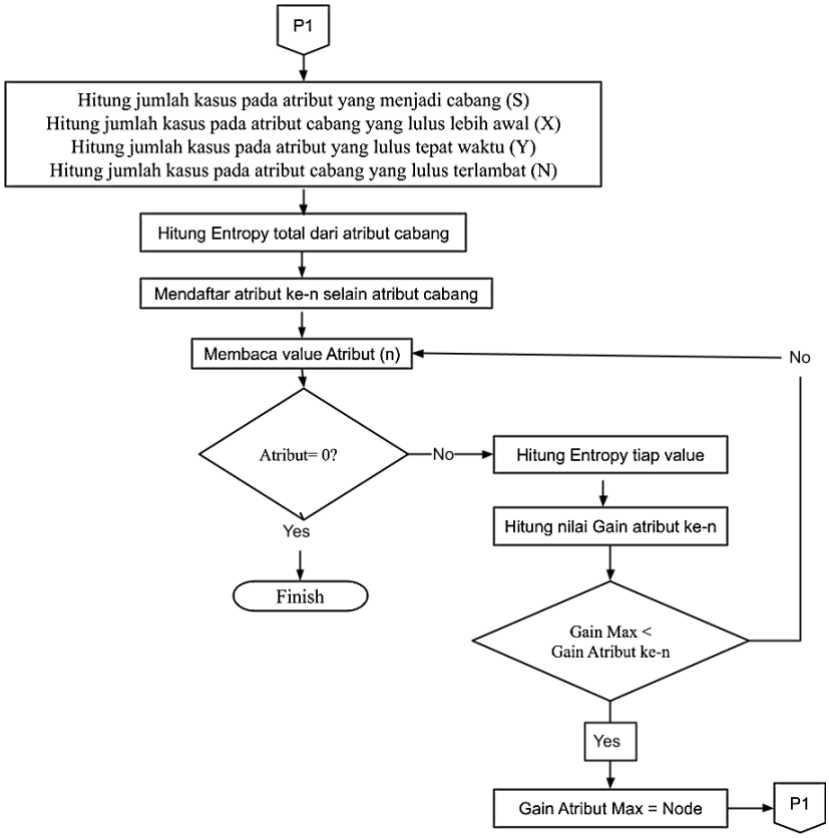
d. Jika nilai *Gain* terbesar telah didapat, sistem akan memilih atribut terbaik untuk dijadikan *root*.
e. Proses ini akan berlangsung sampai semua atribut selesai dihitung.




**Gambar III. 4** *Flowchar*t Penentuan Cabang (Sumber : Data di Olah,2023)

Berikut ini merupakan sistem yang berjalan saat penentuan cabang pada pohon keputusan yang terbentuk (Mashlahah Susi, 2013):

a. Sistem akan memilih atribut dengan nilai *Gain* tertinggi.
b. Nilai pada atribut tertinggi akan diklasifikasikan menurut target yang akan dicapai yaitu lulus lebih awal (X), lulus tepat waktu (Y), lulus terlambat (N).
c. Setiap nilai atribut akan dihitung *Entropy* masing-masing hingga semua atribut selesai dihitung.
d. Jilai nilai *Entropy* = 0 (nol) akan dikoreksi untuk penentuan klasifikasi kasus.
*e.* Jika nilai *Entropy* > 0 (nol) maka akan dijadikan cabang pada *node* selanjutnya.




**Gambar III. 5** *Flowchart* Penentuan *Node* (Sumber : Data di Olah, 2023)

Berikut ini merupakan sistem yang berjalan saat penentuan node pada pembentukan pohon keputusan (Mashlahah Susi, 2013) :

a. Sistem menghitung jumlah kasus pada atribut dengan *Gain* tertinggi
b. Menghitung nilai *Entropy* total dari atribut dengan nilai *Gain* tertinggi.
c. Mendaftar atribut A selain atribut dari nilai *Gain* tertinggi.
d. Menghitung *Entropy* dari setiap nilai atribut ke-n hingga semua terhitung
e. Menghitung nilai *Gain* dari setiap atribut
f. Menentukan *Gain* maksimal sebagai penentuan node selanjutnya
g. Setelah *node* terpilih , proses perhitungan akan berulang dari penentuan cabang, penentuan node hingga semua atribut selesai dieksekusi dan telah mencapai *end of tree.*

## III.5 Perancangan Sistem

Perancangan sistem merupakan uraian yang menggambarkan secara rinci tentang sistem yang akan dibangun. Berikut ini adalah rancangan sistem yang meliputi desain *input*, desain proses, desain *output*, *Use Case Diagram, Activity Diagram, Sequence Diagram*, dan *Class Diagram*.

### *III.5.1* Desain *Input*

Dibawah ini merupakan tabel desain *input* pada sistem informasi prediksi kelulusan mahasiswa Fakultas Sains dan Teknologi UIN Ar-Raniry.

**Tabel III. 2** Desain Input

| **Nama *Input*** | ***Tools Entry* Data** | **Bentuk *Input*** | **Sumber Data** | ***Entry* Data** | **Periode *Input*** | **Deskripsi *Input*** | **Data yang di*entry*** |
|---|---|---|---|---|---|---|---|
| Data Mahasiswa | *Keyboard* dan *Mouse* | Data | Jurusan | Admin | Fleksibel | Berisi data atribut yang digunakan untuk menampilkan data mahasiswa Angkatan 2019 dan 2020 dari setiap jurusan. | Nim, nama, prodi, nilai IP (I-V), IPK terakhir, dan keterangan perkuliahan |
| Data *Training* | *Keyboard* dan *Mouse* | Data | Jurusan | Admin | Fleksibel | Berisi data atribut yang dignakan untuk proses perhitungan | Nim, nama, prodi, IPK, IP (I-V), dan keterangan |
| Data Target | *Keyboard* dan *Mouse* | Data | Jurusan | Admin | Fleksibel | Berisi data atribut yang akan diprediksi kelulusannya . | Nim, nama, prodi, IPK, IP (I-V), dan keterangan |

*III.5.2* Desain *Output*

Dibawah ini merupakan tabel desain *output* pada sistem informasi prediksi kelulusan mahasiswa Fakultas Sains dan Teknologi UIN Ar-Raniry yang menampilkan hasil prediksi dan ditampilkan dengan *chart*.

**Tabel III. 3** Desain *Output*

| No | Nama Laporan | Bentuk | Periode | *Tools* | Data yang ditampilkan | Deskripsi |
|---|---|---|---|---|---|---|
| 1 | Prediksi Kelulusan | Tabel pada form | Fleksibel | Monitor | Prediksi kelulusan mahasiswa | Berisi data hasil prediksi |
| 2 | *Chart* kelulusan | *Chart* pada form | Fleksibel | Monitor | *Chart* jumlah mahasiswa yang lulus lebih awal, tepat waktu, dan terlambat | Tampilan *chart* kelulusan mahasiswa |

### III.5.3 Desain Proses

Dibawah ini merupakan tabel desain proses pada sistem informasi prediksi kelulusan mahasiswa Fakultas Sains dan Teknologi UIN Ar-Raniry yang terdiri dari *input* data, hapus data, tampil data, serta menghitung nilai *Entropy* dan *Gain*.

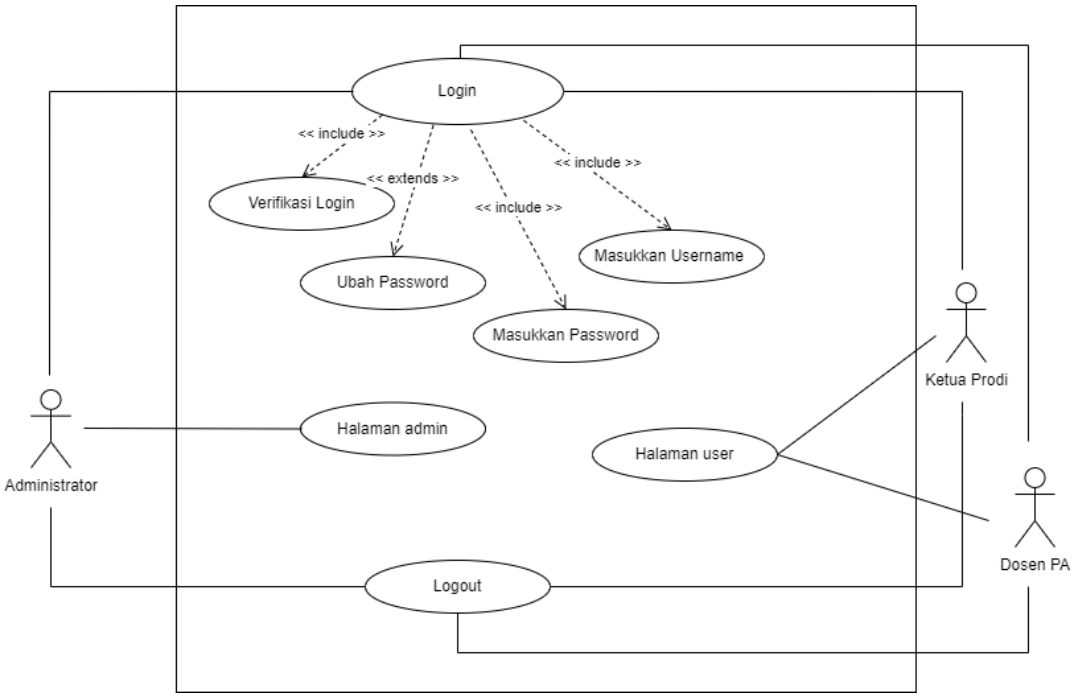
**Tabel III. 4** Desain Proses

| Nama Proses | Deskripsi Proses | Proses | Output Proses |
|---|---|---|---|
| *Input* Data | Memasukkan data | Memasukkan data | *Database* |
| Hapus data dan tampil data | Menampilkan dan menghapus untuk melakukan *update* | Menghapus dan menampilkan | Tabel berisi tampilan data dari *database* |
| Hitung *Entropy* dan *Gain* | Mengambil data *training* untuk diklasifikasikan menurut atribut dan menghitung nilai *Entropy* dan *Gain* menggunakan algoritma C4.5 | Klasifikasi / Penentuan *Rule* | *Rule* atau aturan yang dihasilkan |

III.5.4 *Use Case* Diagram

*Use Case* diagram pada sistem informasi prediksi kelulusan hanya terdiri dari tiga aktor yaitu Admin, Ketua Prodi, dan Dosen PA (Penasehat Akademik). Admin merupakan aktor yang mempunyai hak askses untuk login dan bertanggung untuk mengelola keseluruhan sistem dari master data yang meliputi data prodi dan mahasiswa, hingga melakukan prediksi serta mencetak hasil prediksi.

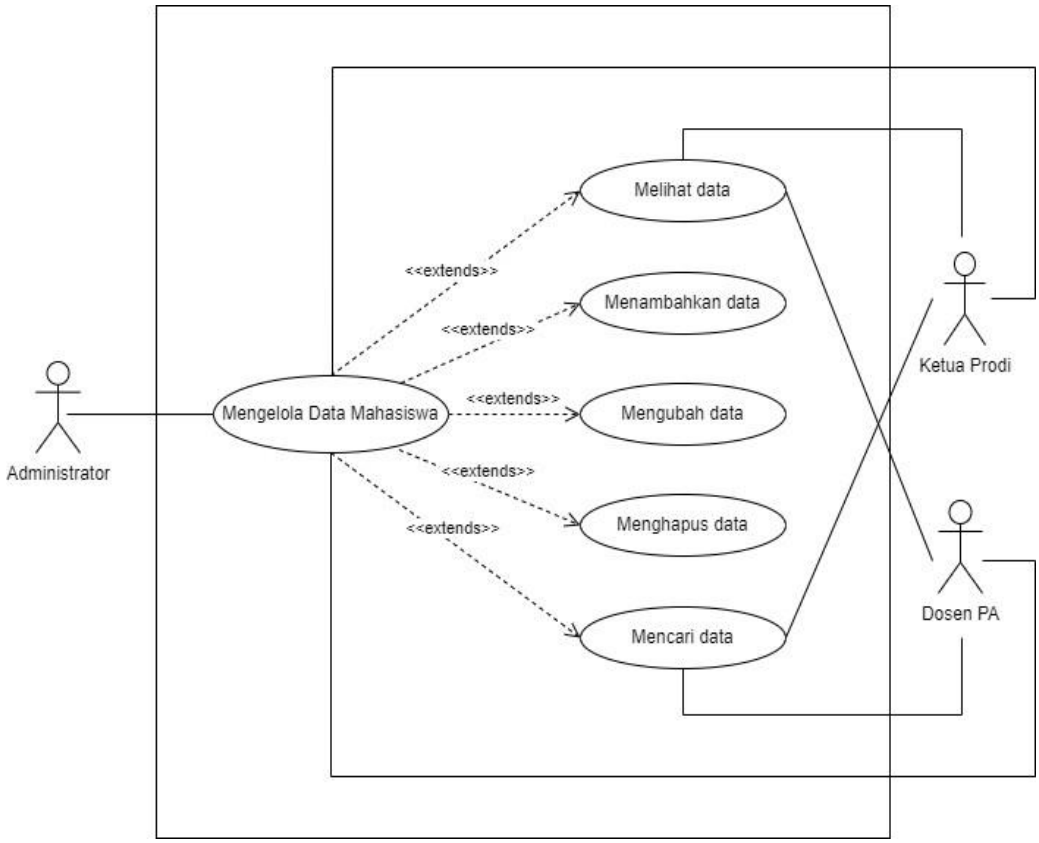
1) *Use Case Login*, Ubah Password, *Logout*




**Gambar III. 6** *Use Case Login*

Pada Gambar III.6 merupakan *use case login*, ubah *password*, dan *logout* sistem yang dilakukan oleh tiga aktor yaitu admin, ketua prodi, dan dosen PA. Admin memiliki tampilan halaman yang berbeda dengan ketua prodi dan dosen PA.

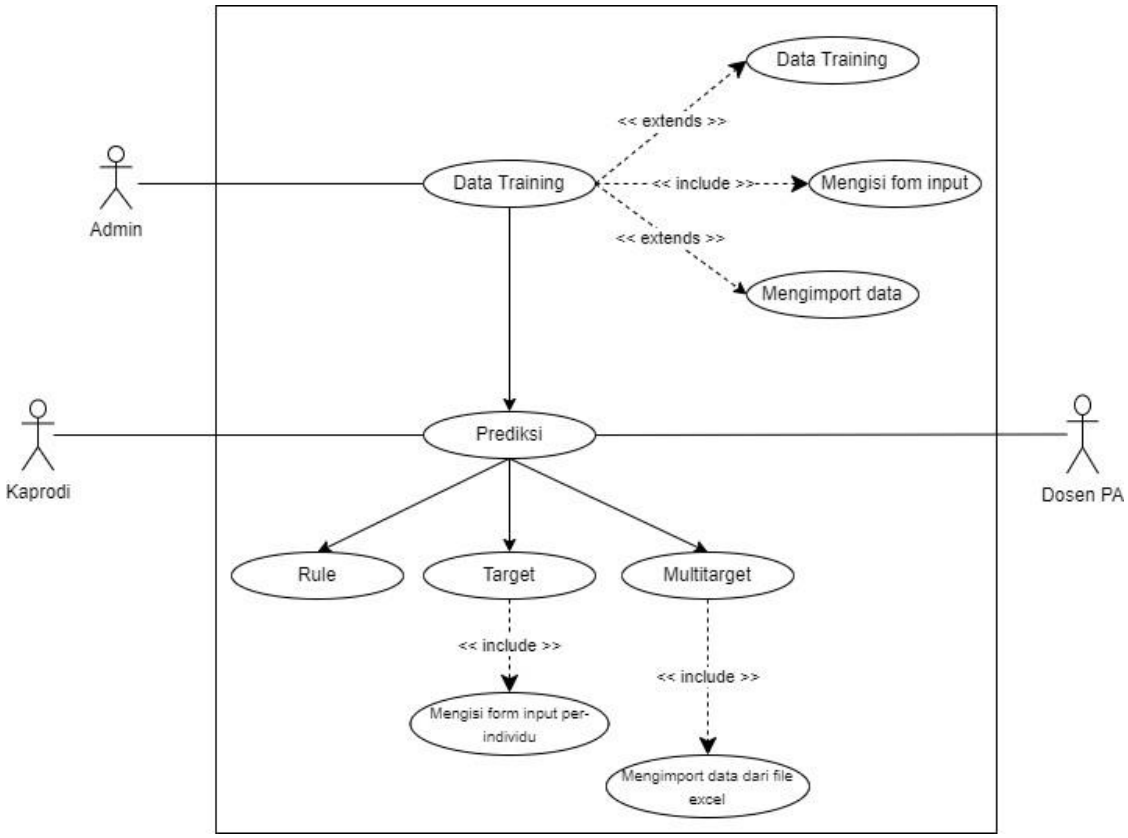
2) *Use Case* Data Mahasiswa




**Gambar III. 7** *Use Case* Data Mahasiswa

Pada Gambar III.7 merupakan *use case* data mahasiswa dari tiga aktor yaitu admin, ketua prodi, dan dosen PA. Admin mempunyai hak akses untuk mengelola keseluruhan data mahasiswa seperti melihat data, menambahkan data, mengubah data, menghapus data dan melakukan pencarian data. Sedangkan ketua prodi dan dosen PA hanya memiliki akses untuk melihat dan melakukan pencarian data.

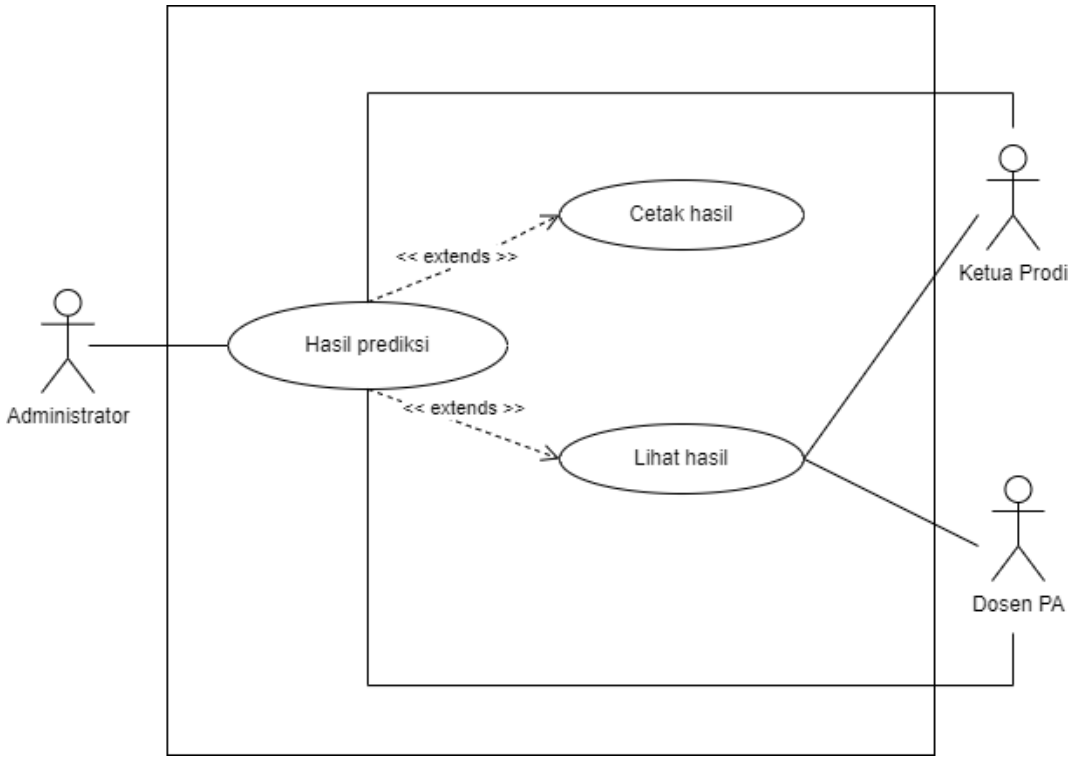
3) *Use Case* Prediksi




**Gambar III. 8** *Use Case* Prediksi

Pada Gambar III.8 merupakan *use case* prediksi dari tiga aktor yaitu admin, ketua prodi, dan dosen PA. Admin mempunyai hak akses untuk melakukan data *training*, *rule*, target, dan multitarget. Sedangkan aktor Kaprodi dan Dosen PA hanya memiliki akses untuk melakukan *rule*, target, dan multi target.

4) *Use Case* Laporan Hasil

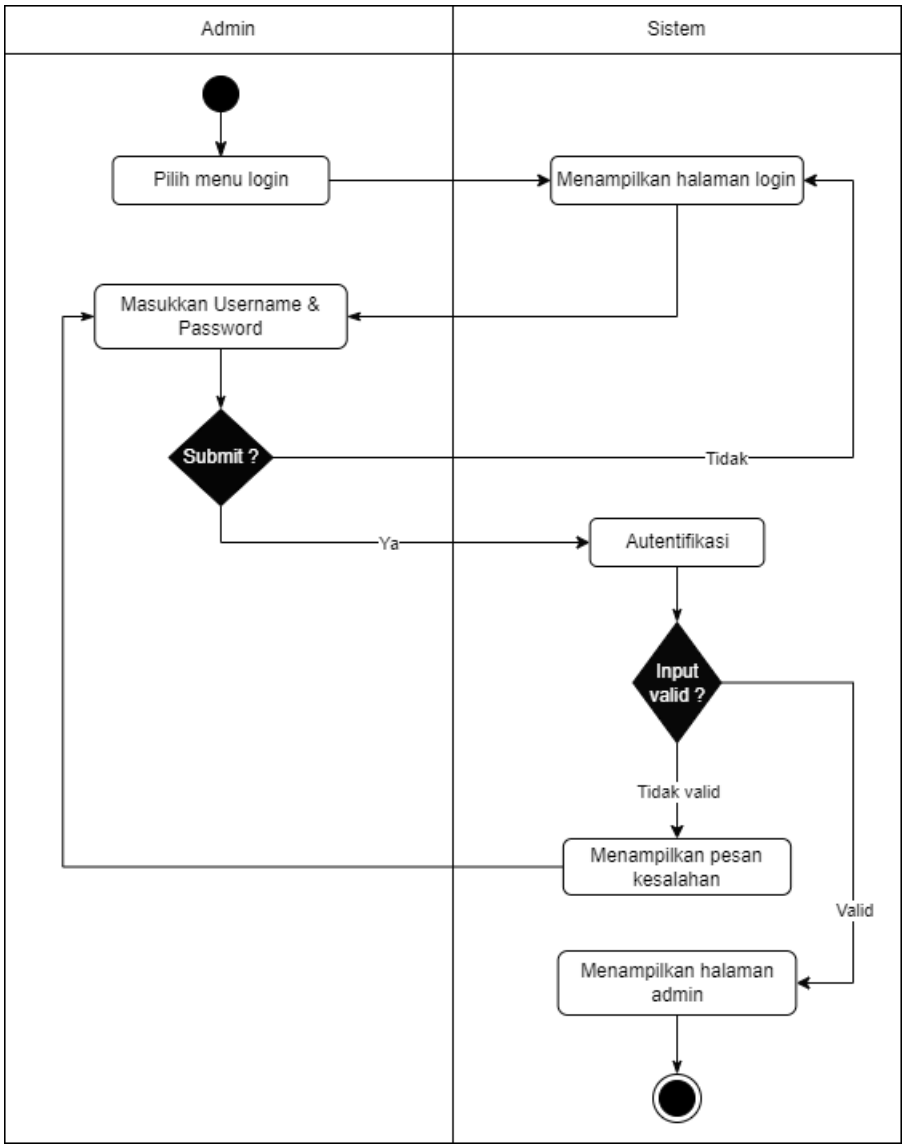



**Gambar III. 9** *Use Case* Laporan Hasil

Pada Gambar III.9 merupakan *use case* laporan hasil prediksi dari tiga aktor yaitu admin, ketua prodi, dan dosen PA. Admin mempunyai hak akses untuk mengelola keseluruhan hasil prediksi. Sedangkan ketua prodi dan dosen PA hanya memiliki akses untuk melihat hasil prediksi.

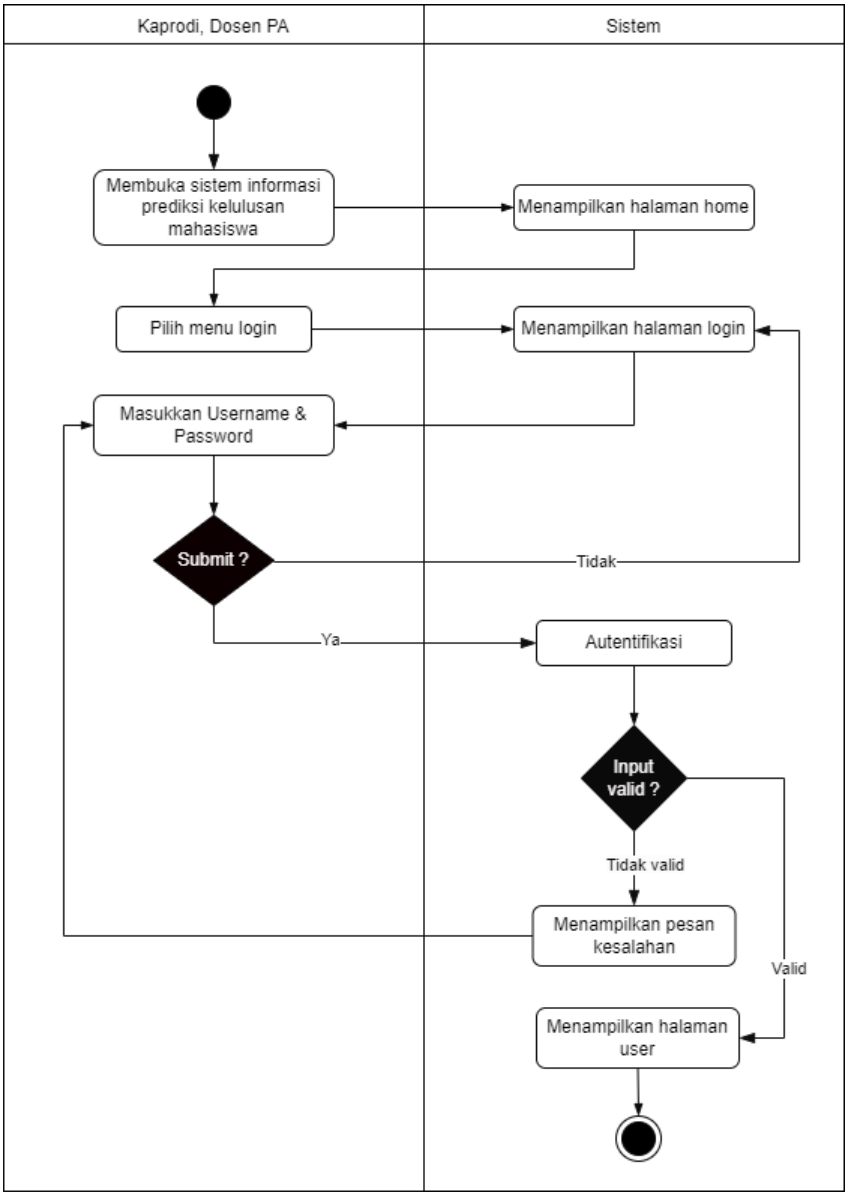
*III.5.5 Activity Diagram*

*Activity diagram* pada sistem informasi prediksi kelulusan terdiri dari 11 gambar aktivitas yang dilakukan oleh Admin, Kaprodi, Dosen PA. Admin merupakan aktor yang mempunyai hak akses untuk bertanggung untuk mengelola keseluruhan sistem dari master data yang meliputi data prodi dan mahasiswa, hingga melakukan prediksi serta mencetak hasil prediksi. Sedangkan Kaprodi, dan Dosen PA hanya dapat melakukan *login, logout*, ubah password, prediksi, lihat data prodi, lihat data mahasiswa, dan lihat hasil prediksi.



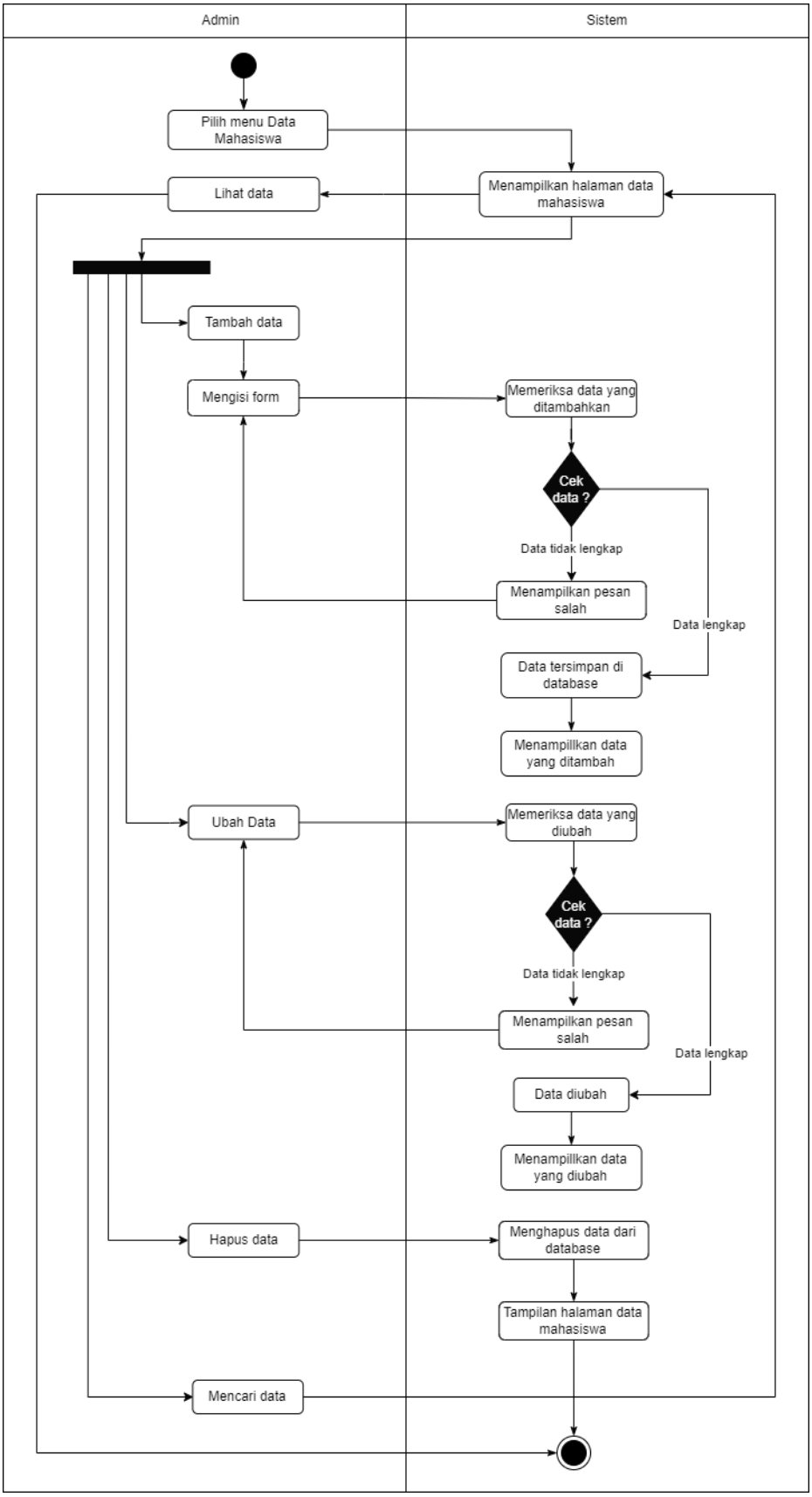

**Gambar III. 10** *Activity* Admin *Login*

Pada Gambar III.10 menjelaskan proses admin untuk melakukan aktivitas *login* dengan memasukkan *username* dan *password*. Jika data sesuai maka admin dapat masuk ke halaman admin. Apabila gagal maka admin akan kembali ke halaman *login*.



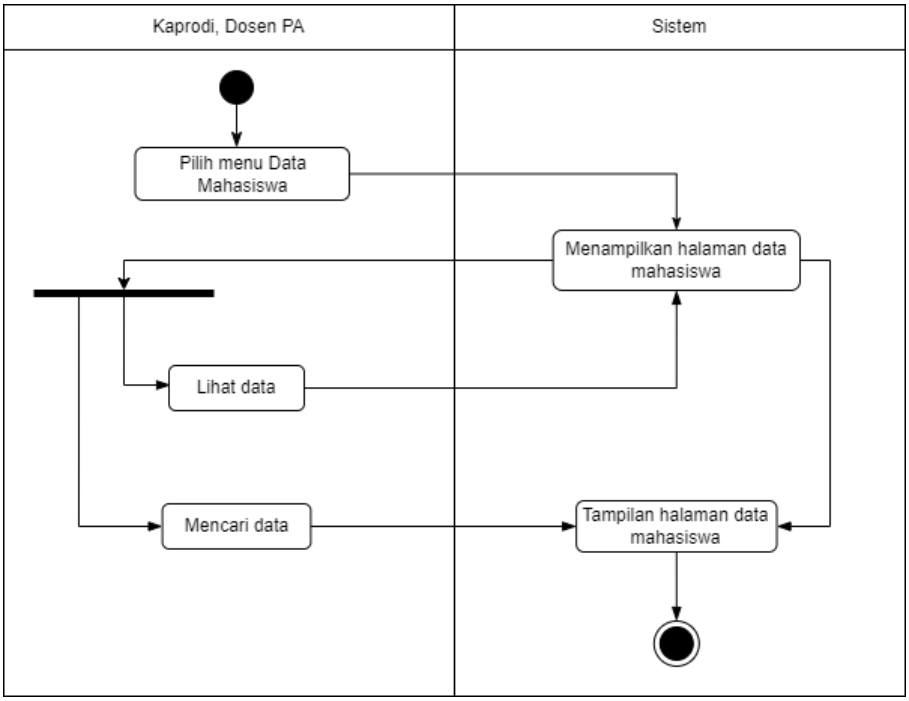

**Gambar III. 11** *Activity User Login*

Pada Gambar III.11 menjelaskan proses *user* Kaprodi dan Dosen PA untuk melakukan aktivitas *login* dengan memasukkan *username* dan *password*. Jika data sesuai maka *user* dapat masuk ke halaman *user*. Apabila gagal maka *user* akan kembali ke halaman *login*.




**Gambar III. 12** *Activity* Admin pada data mahasiswa

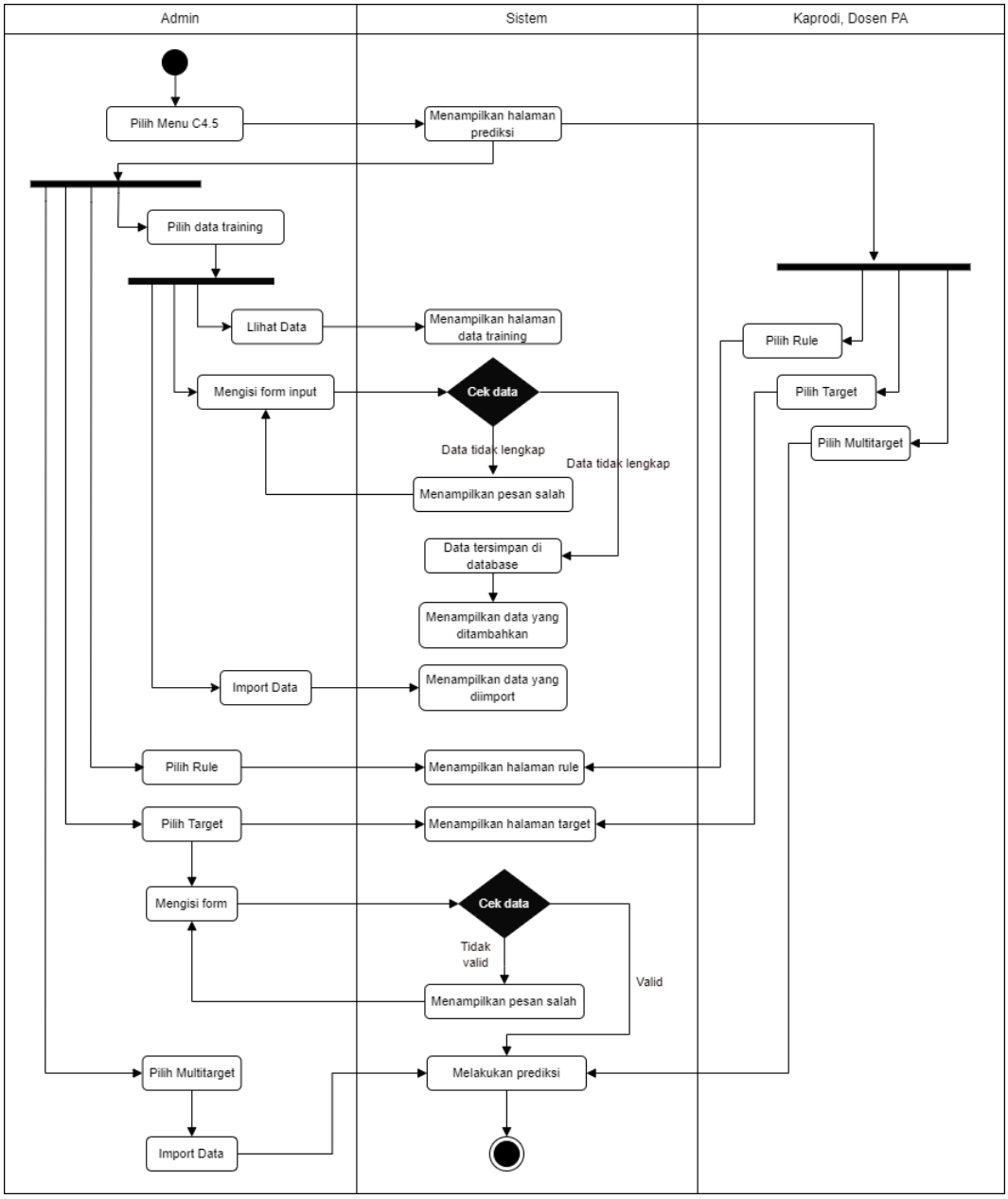
Pada Gambar III.12 menjelaskan proses aktivitas admin pada halaman data mahasiswa. Admin memiliki hak akses untuk melihat data, menambah data, mengubah data, menghapus data, dan melakukan pencarian data mahasiswa. Proses menambah data dapat dilakukan dengan mengisi *form* yang kemudian akan dilakukan pengecakan data oleh sistem. Jika data yang ditambahkan sesuai maka data akan dikirim ke *database*, tetapi jika belum sesuai maka admin akan kembali ke form *input* data. Untuk proses mengubah dan menghapus data, dapat dilakukan admin dengan meng-klik ikon tersebut. Kemudian sistem akan memproses data yang akan dihapus atau yang diubah dan menampilkan halaman data mahasiswa yang telah diperbarui.




**Gambar III. 13** *Activity User* pada Data Mahasiswa

Pada Gambar III.13 menjelaskan proses aktivitas *user* Kaprodi dan Dosen PA pada halaman data mahasiswa. *User* hanya memiliki hak akses untuk melihat data, dan melakukan pencarian data mahasiswa. Proses pencarian data dapat dilakukan dengan mengetikkan data yang ingin dicari pada kolom *search*,

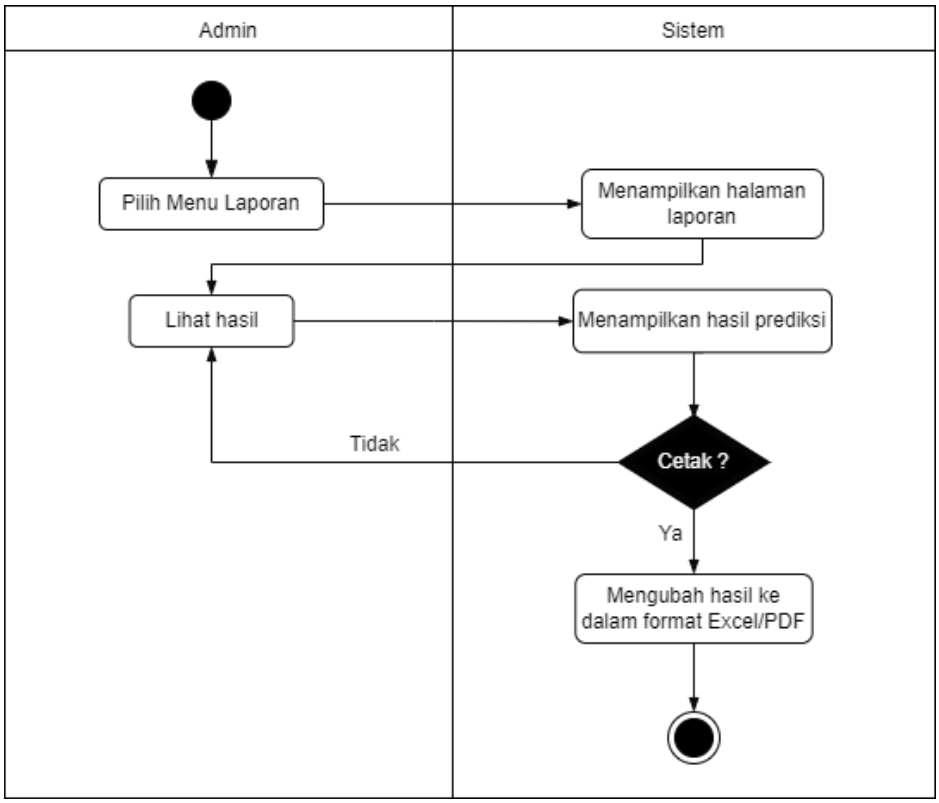
kemudian akan dilakukan pengecakan data oleh sistem dan menampilkan data tersebut.




**Gambar III. 14** *Activity* Prediksi

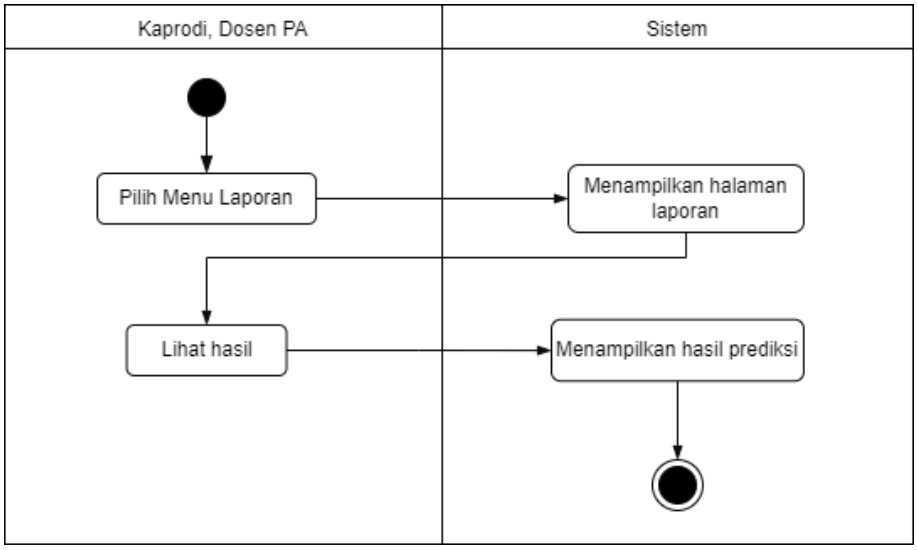
Pada Gambar III.14 menjelaskan proses aktivitas Admin, Kaprodi, dan Dosen PA pada halaman prediksi. Ketiga aktor memiliki hak akses untuk melakukan prediksi kelulusan mahasiswa yang terdiri dari data *training*, *rule*, dan target, serta multitarget untuk melakukan prediksi dengan menggunakan banyak

data sekaligus. Pada data *training*, data yang akan dimasukkan dapat dilakukan dengan mengisi form atau dengan meng-*import* data. Data yang telah diprediksi akan ditampilkan oleh sistem jika semua proses telah dilakukan.



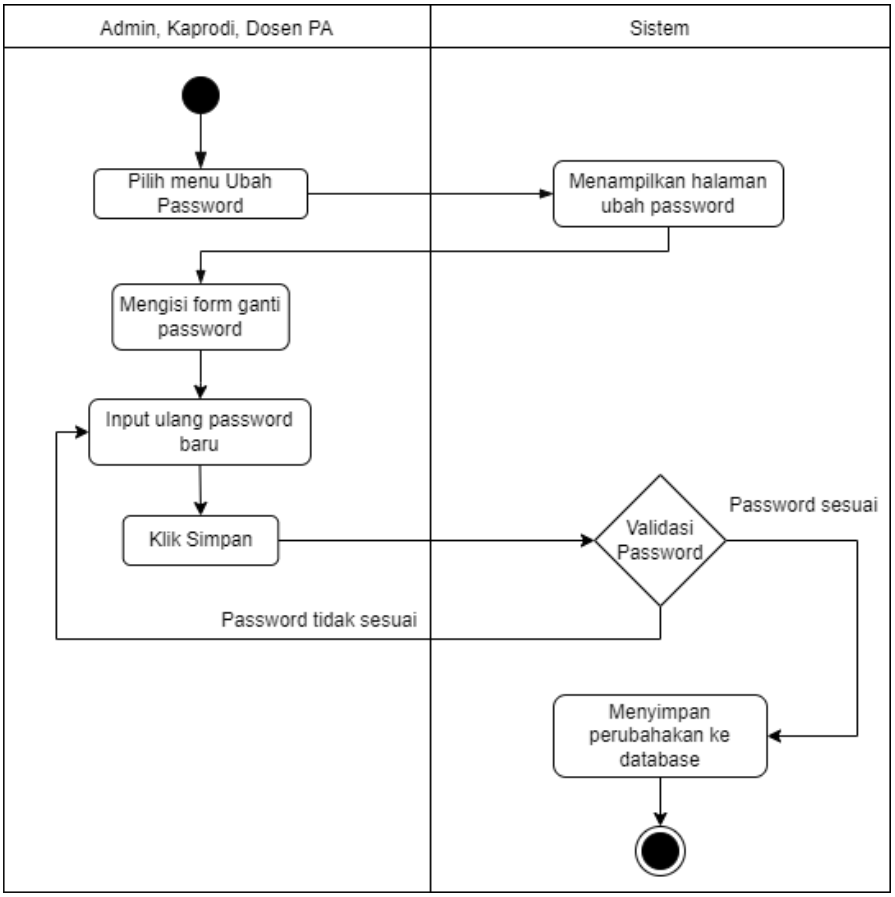

**Gambar III. 15** *Activity* Laporan Admin

Pada Gambar III.15 menjelaskan proses aktivitas admin pada halaman laporan. Admin memiliki hak akses untuk melihat data, dan mencetak hasil. Proses cetak data tersebut akan dilakukan oleh sistem dengan mengubah format data dalam bentuk *Excel* atau PDF.



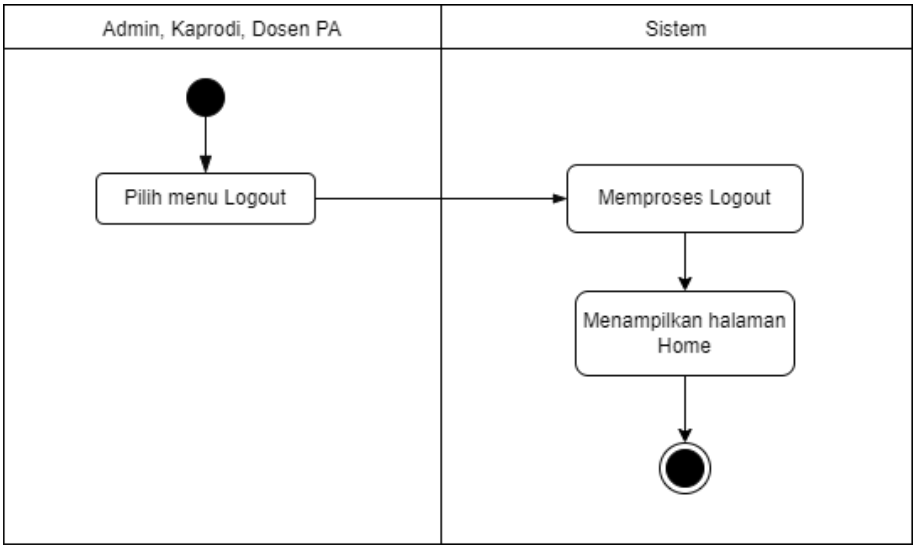

**Gambar III. 16** *Activity* Laporan *User*

Pada Gambar III.16 menjelaskan proses aktivitas *user* Kaprodi dan Dosen PA pada halaman laporan. *User* hanya memiliki hak akses untuk melihat hasil prediksi yang ditampilkan oleh sistem.




**Gambar III. 17** *Activity* Ubah password

Pada Gambar 3.17 menjelaskan proses aktivitas Admin, Kaprodi dan Dosen PA pada halaman ubah *password*. Ketiga aktor memiliki hak akses untuk mengubah *password* dengan mengisi form ganti *password* dan mengulangi *password* baru yang telah diubah. Kemudian sistem akan memvalidasi *password* baru. Jika *password* tidak sesuai maka akan kembali ke bagian *input* ulang *password* baru, tetapi jika *password* sudah sesuai akan disimpan perubahan tersebut ke *database*.

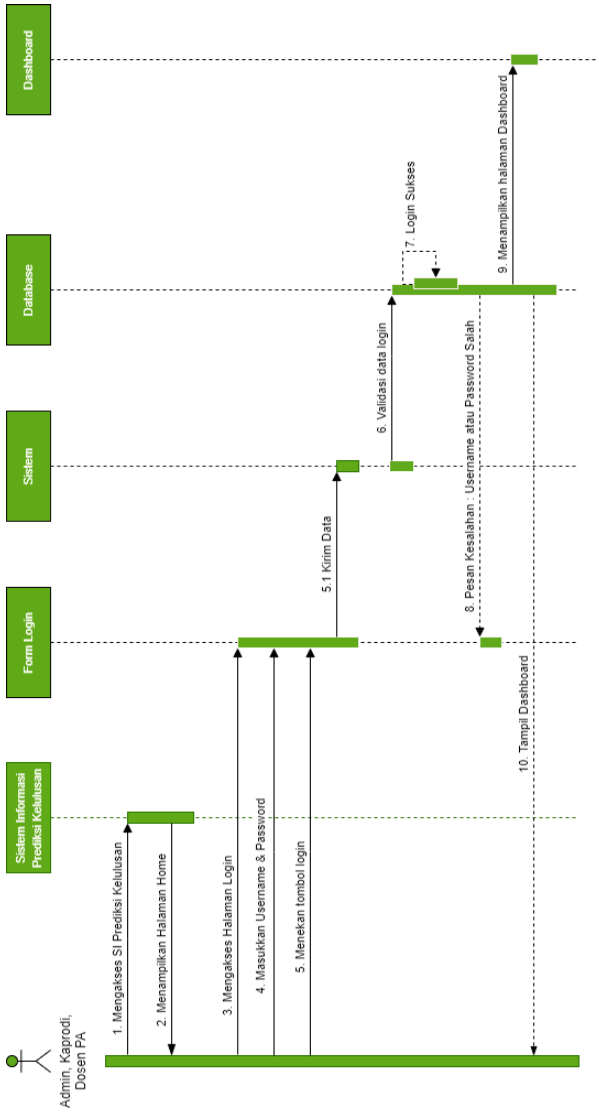



**Gambar III**. **18** *Activity Logout*

Pada Gambar III.18 menjelaskan proses aktivitas Admin, Kaprodi dan Dosen PA pada halaman *logout*. Sistem akan memproses dan menampillkan kembali halaman *home.*

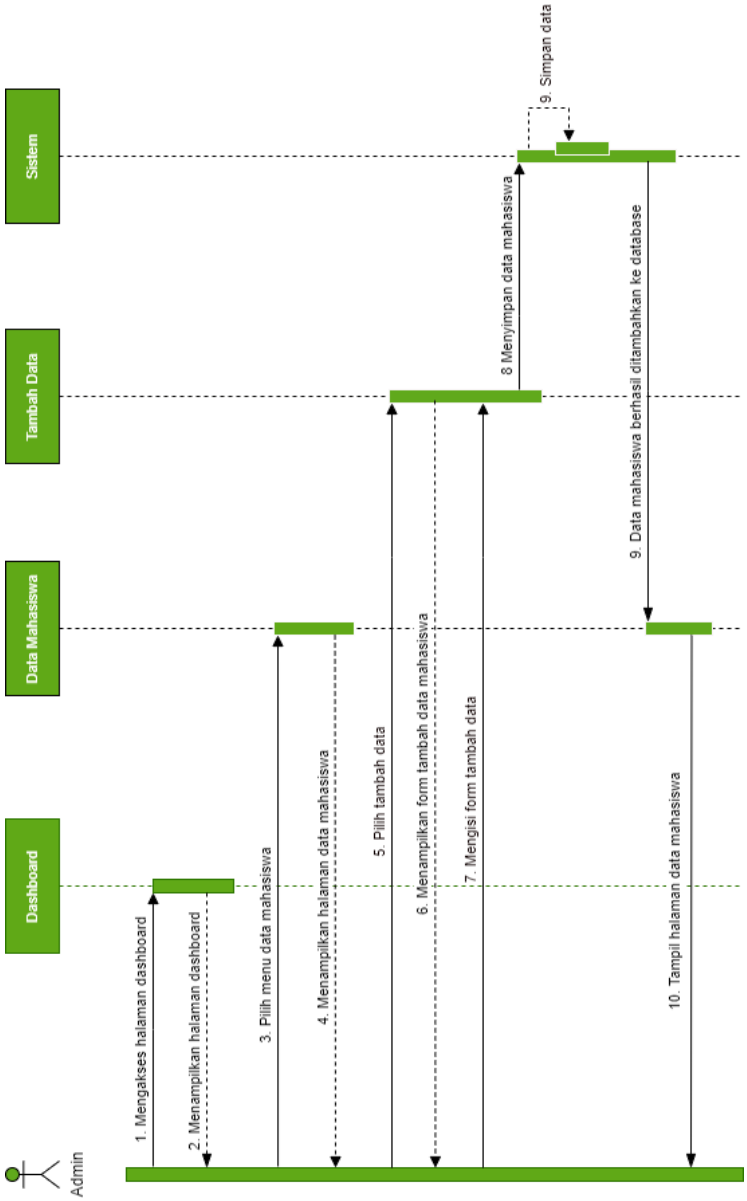
*III.5.6 Sequence Diagram*

*Sequence diagram* merupakan *grafis* yang menggambarkan interaksi objek melalui pesan pada eksekusi *use case*. Berikut gambar *sequence diagram* "Sistem Informasi Prediksi Kelulusan Mahasiswa Fakultas Sains dan Teknologi" :



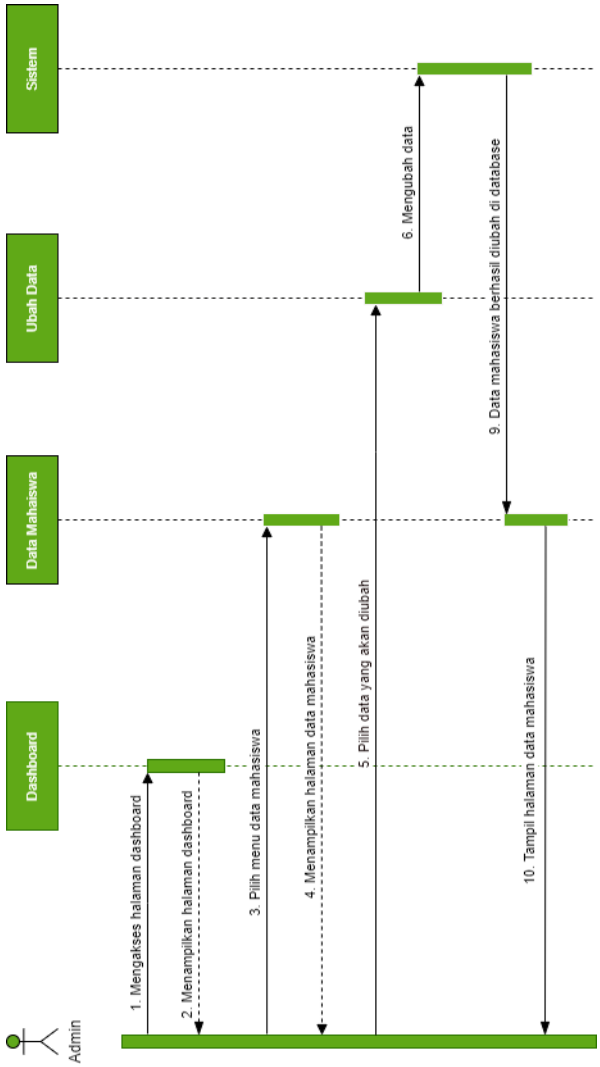

**Gambar III. 19** *Sequence Login*

Pada Gambar III.19 menjelaskan proses *sequence login* yang dilakukan oleh Admin, Kaprodi, dan Dosen PA dengan mengakses sistem informasi prediksi kelulusan, memasukkan *username* dan *password*. Setelah itu, data akan dikirim dan di validasi oleh sistem. Jika tervalidasi, maka admin akan diarahkan ke halaman admin, kaprodi, dan dosen PA ke halaman *user*.



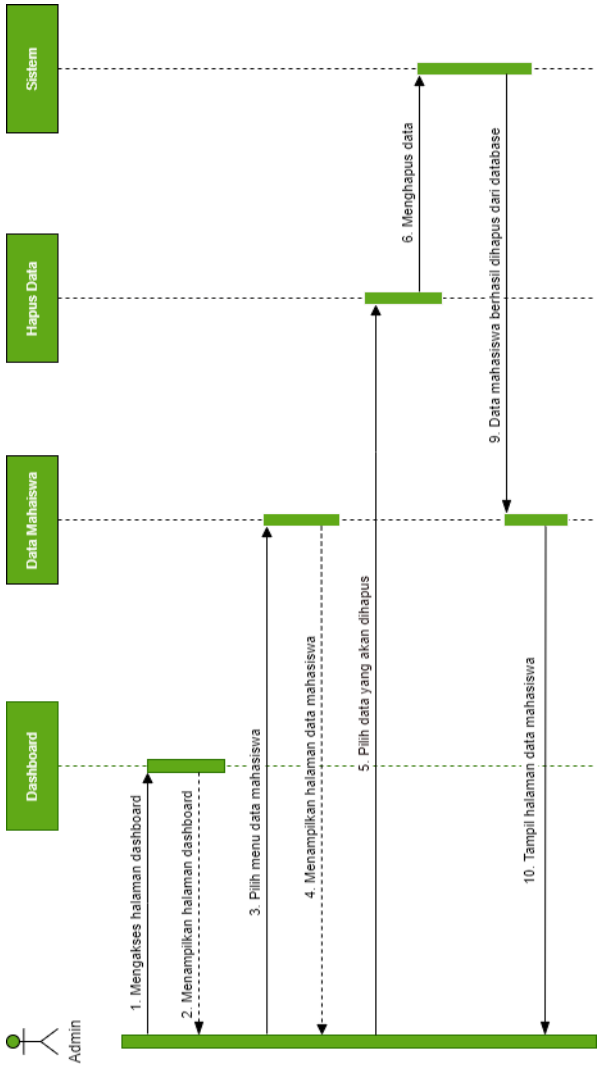

**Gambar III. 20** *Sequence* Admin Tambah Data Mahasiswa

Pada Gambar III.20 menjelaskan proses *sequence* tambah data mahasiswa yang dilakukan oleh Admin dengan mengisi form tambah data. Setelah itu, data akan dikirim dan di simpan oleh sistem ke dalam *database*. Jika data berhasil tersimpan ke dalam *database*, maka sistem akan menampilkan kembali data yang telah ditambahkan ke halaman data mahasiswa.



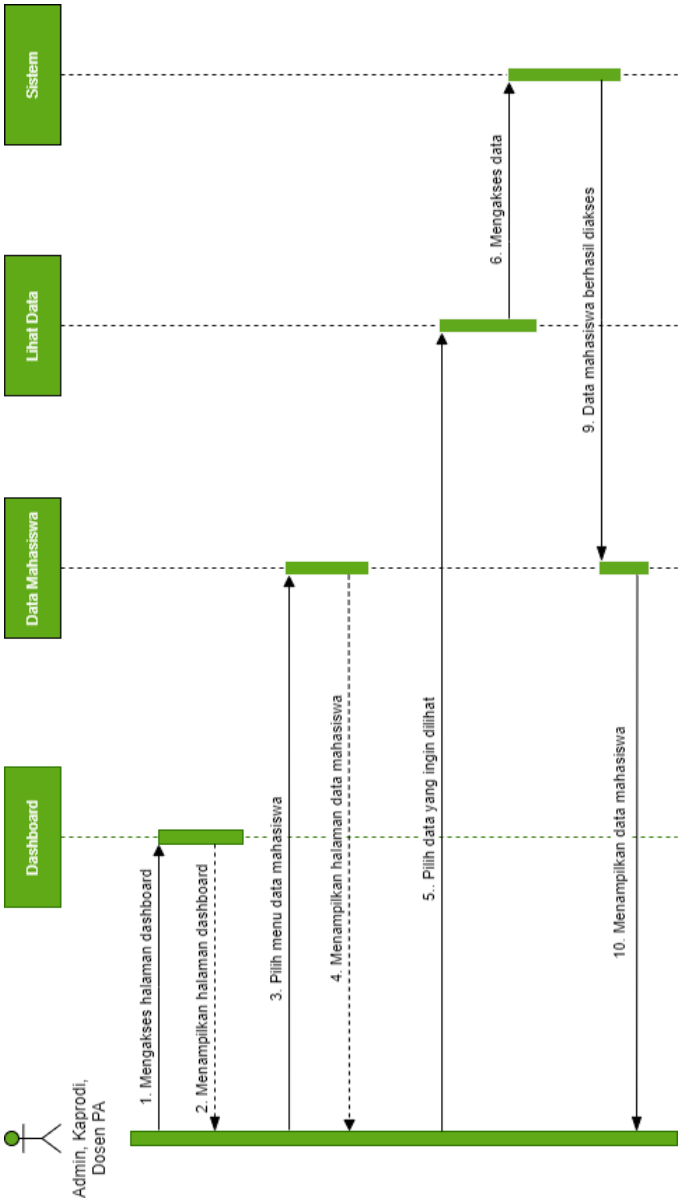

**Gambar III. 21** *Sequence* Admin Ubah Data Mahasiswa

Pada Gambar III.21 menjelaskan proses *sequence* ubah data mahasiswa yang dilakukan oleh admin dengan mengisi memilih data yang akan diubah. Setelah itu, data akan dikirim oleh sistem ke dalam *database*. Jika data berhasil diubah ke dalam *database*, maka sistem akan menampilkan kembali data yang telah diperbarui di halaman data mahasiswa.



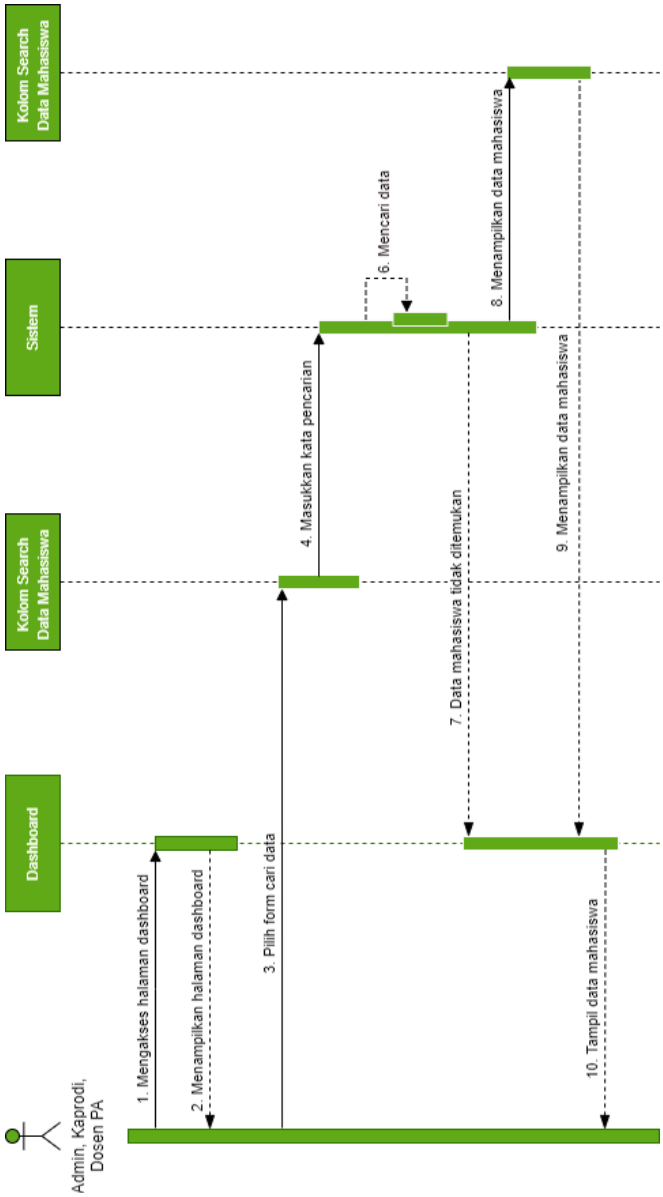

**Gambar III. 22** *Sequence* Admin Hapus Data Mahasiswa

Pada Gambar III.22 menjelaskan proses *sequence* hapus data mahasiswa yang dilakukan oleh admin dengan memilih data yang akan dihapus. Setelah itu, sistem akan menghapus data tersebut dari *database*. Jika data berhasil dihapus, maka sistem akan menampilkan kembali data yang telah diperbarui di halaman data mahasiswa.



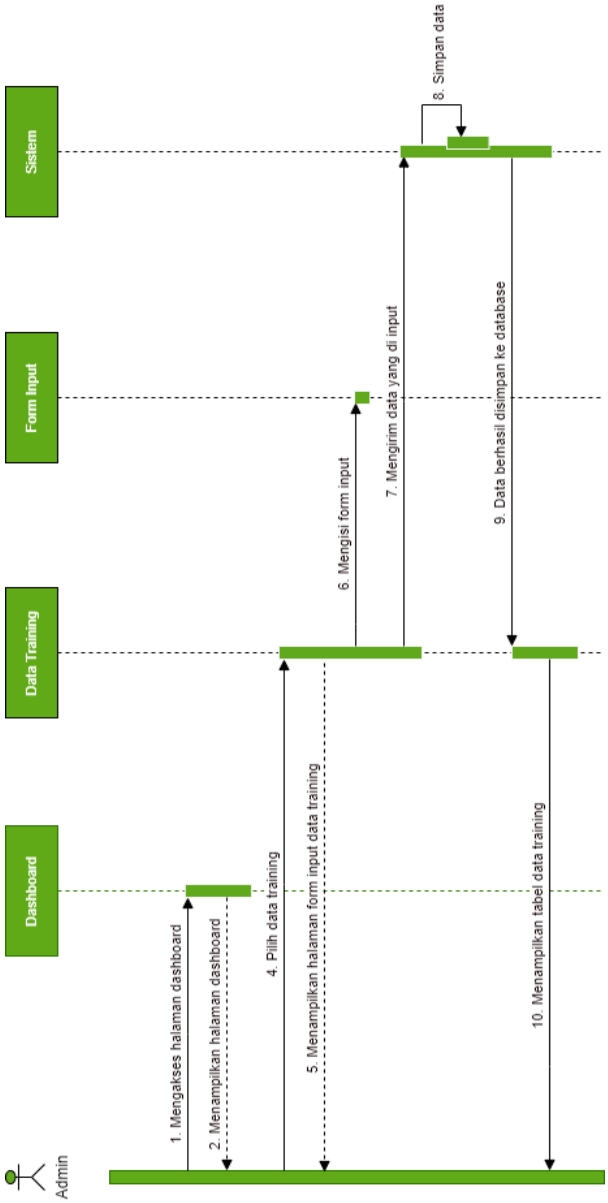

**Gambar III. 23** *Sequence* Lihat Data Mahasiswa

Pada Gambar III.23 menjelaskan proses *sequence* lihat data mahasiswa yang dilakukan oleh Admin, Kaprodi, dan Dosen PA. Sistem akan mengakses data dan menampilkan data tersebut di halaman data mahasiswa.



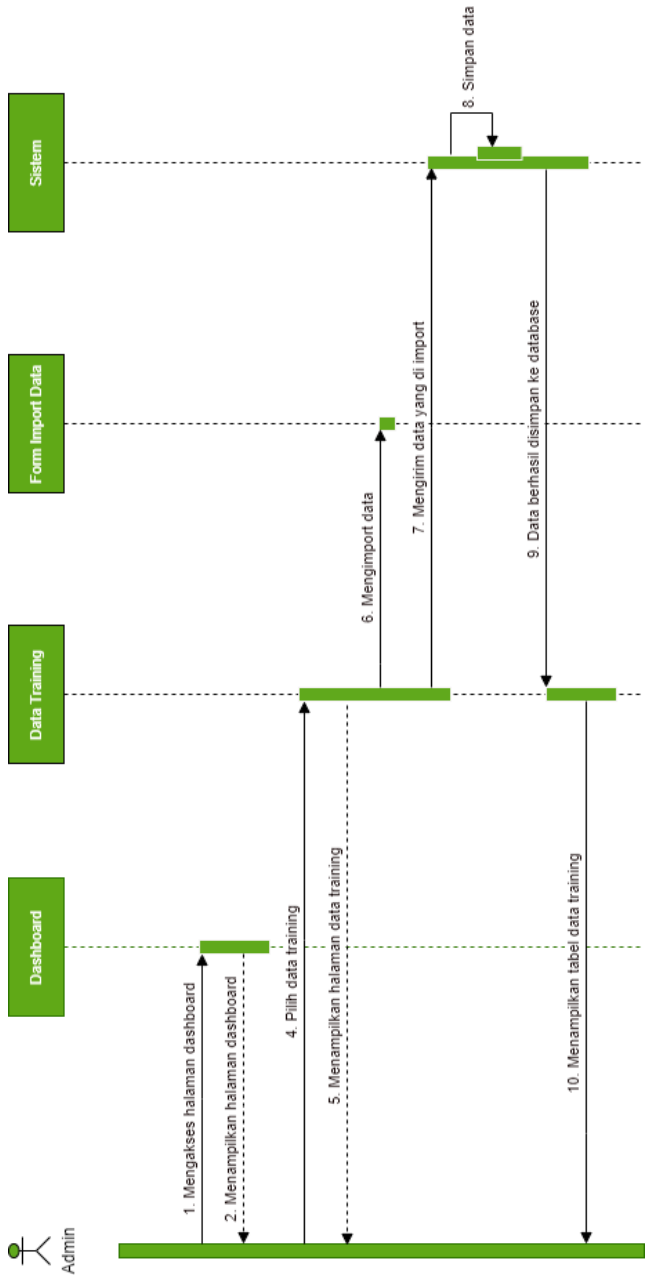

**Gambar III. 24** *Sequence Search* Data Mahasiswa

Pada Gambar III.24 menjelaskan proses *sequence* mencari data mahasiswa yang dilakukan oleh Admin, Kaprodi, dan Dosen PA pada kolom pencarian. Sistem akan mencari data tersebut. Jika data ditemukan, sistem akan menampilkan data dihalaman data mahasiswa, tetapi jika data tidak ditemukan, maka ketiga aktor harus memasukkan kembali data yang ingin dicari.



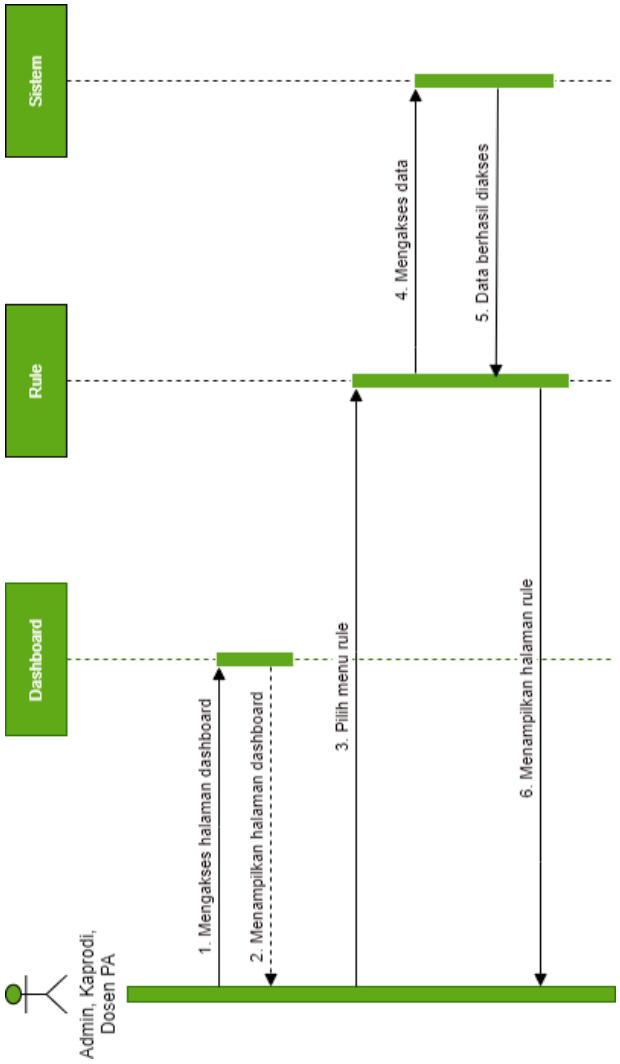

**Gambar III. 25** *Sequence* Admin *Input* Data *Training*

Pada Gambar III.25 menjelaskan proses *sequence input* data *training* yang dilakukan oleh admin dengan mengisi form tersebut. Data akan dikirim dan disimpan oleh sistem ke dalam *database*. Setelah itu, sistem akan menampilkan kembali data yang telah di *input* ke halaman data *training*.



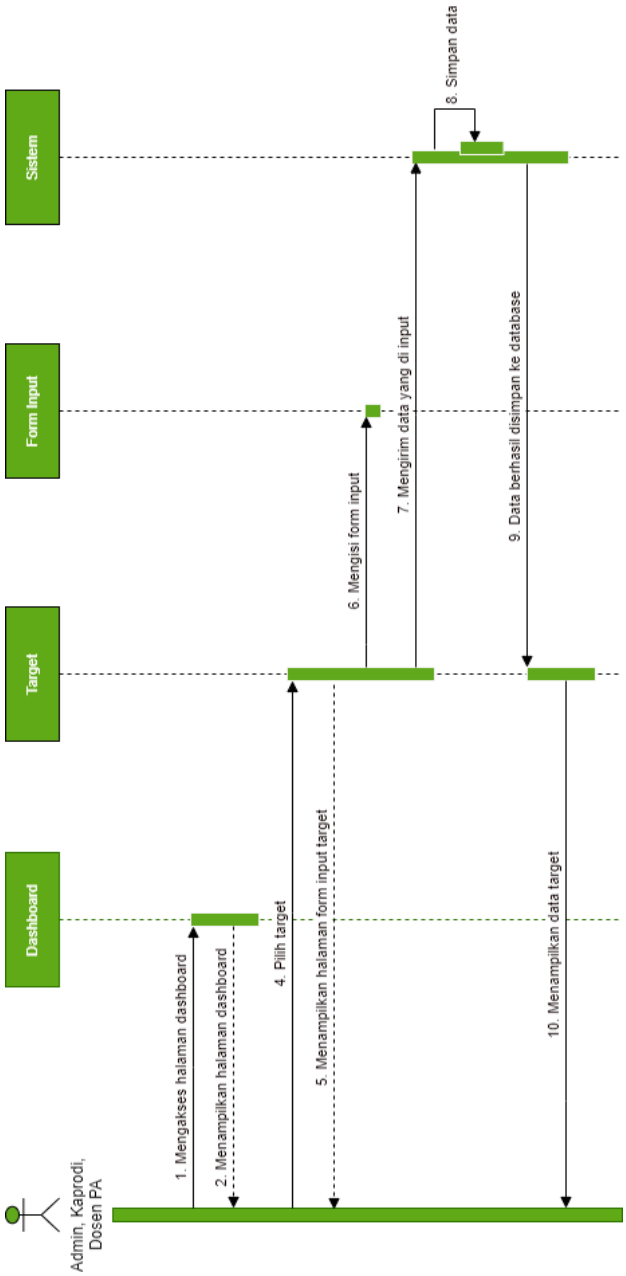

**Gambar III. 26** *Sequence* Admin *Import* Data *Training*

Pada Gambar III.26 menjelaskan proses *sequence import* data *training* yang dilakukan oleh admin. Sistem akan memproses data yang di *import* dan menyimpannya ke dalam *database*. Setelah itu, data yang telah di *import* akan ditampilkan di halaman data *training*.



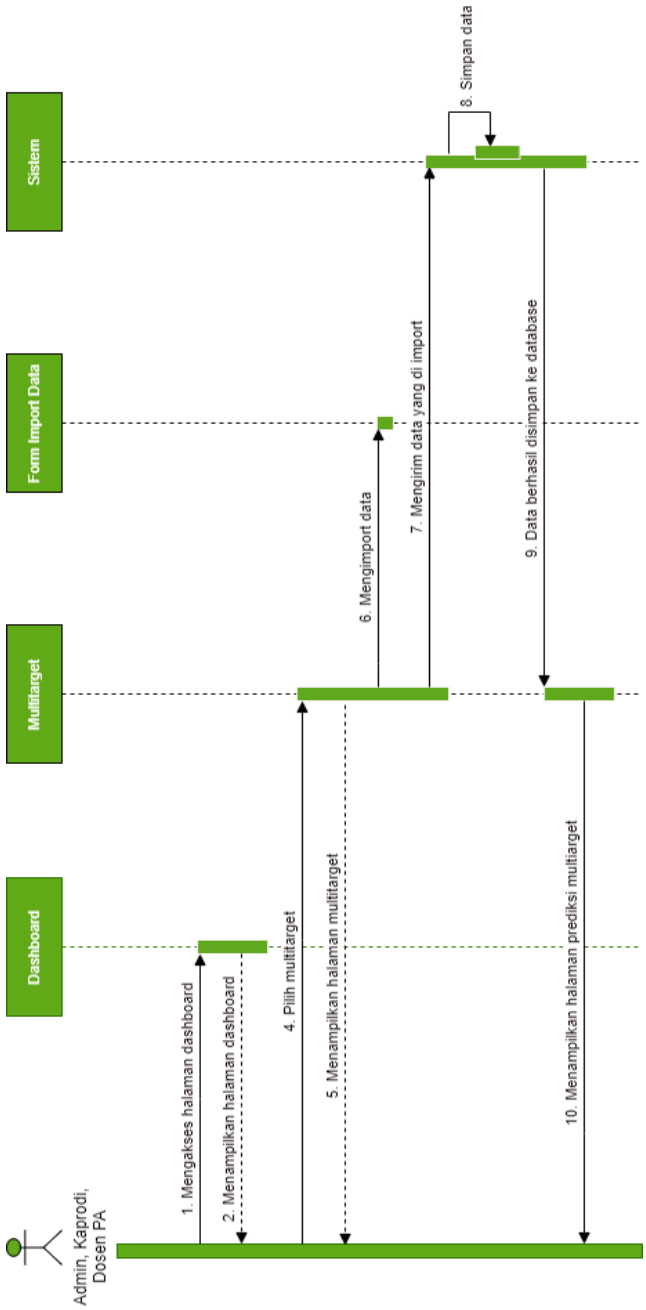

**Gambar III. 27** *Sequence Rule*

Pada Gambar III.27 menjelaskan proses *sequence rule* dilakukan oleh Admin, Kaprodi, dan Dosen PA. Sistem akan mengakses data dan menampilkannya dihalaman *rule.*



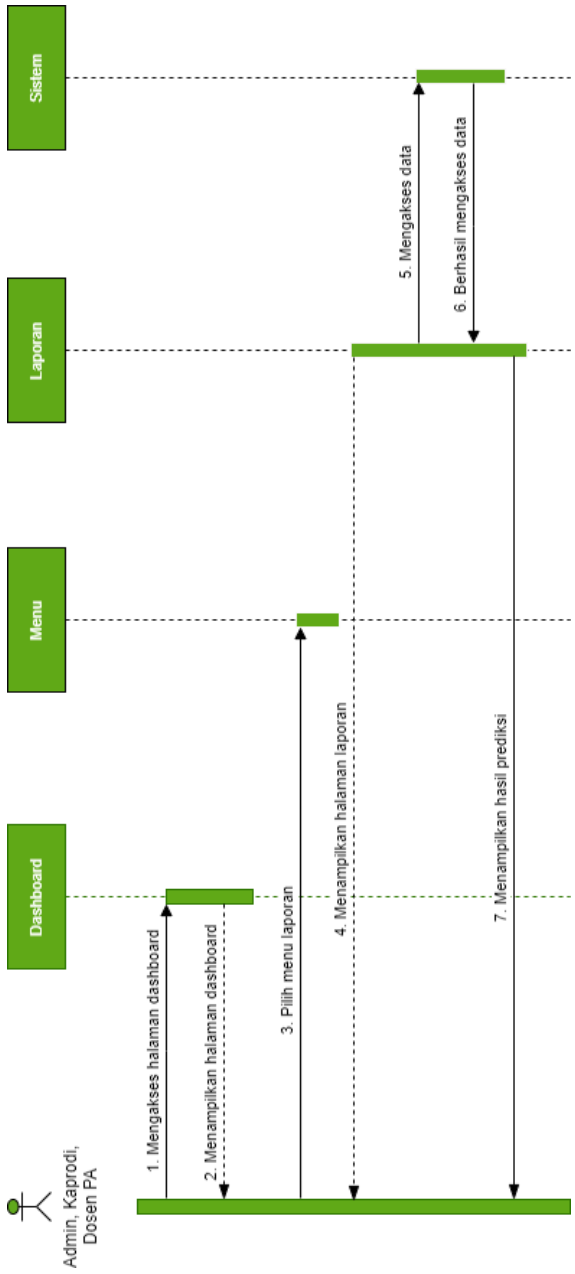

**Gambar III. 28** *Sequence Input* Target

Pada Gambar III.28 menjelaskan proses *sequence* input target yang dilakukan oleh admin dengan mengisi form. Data akan dikirim dan disimpan oleh sistem ke dalam *database*. Setelah itu, sistem akan menampilkan kembali data yang telah di *input* ke halaman target.



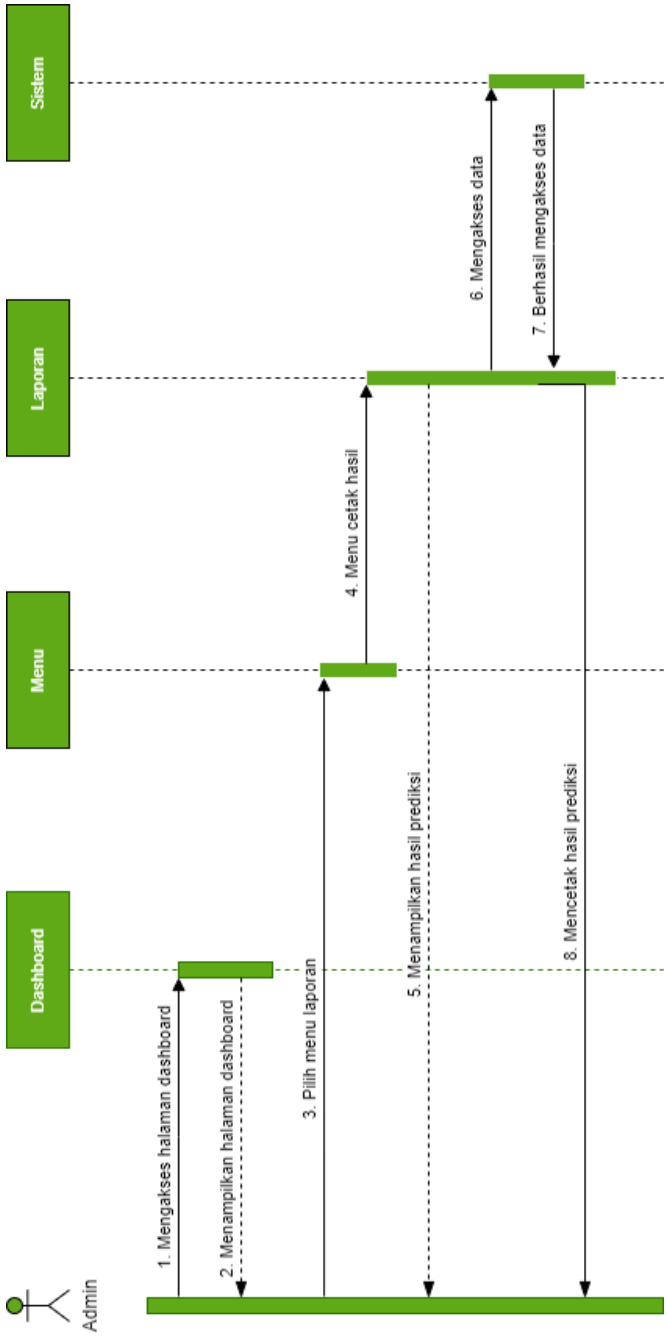

**Gambar III. 29** *Sequence Import* Data Multitarget

Pada Gambar III.29 menjelaskan proses *sequence input* multitarget yang dilakukan oleh admin dengan meng-*import* data dari *excel*. Data akan dikirim dan disimpan oleh sistem ke dalam *database*. Setelah itu, sistem akan menampilkan kembali data dengan hasil prediksi.



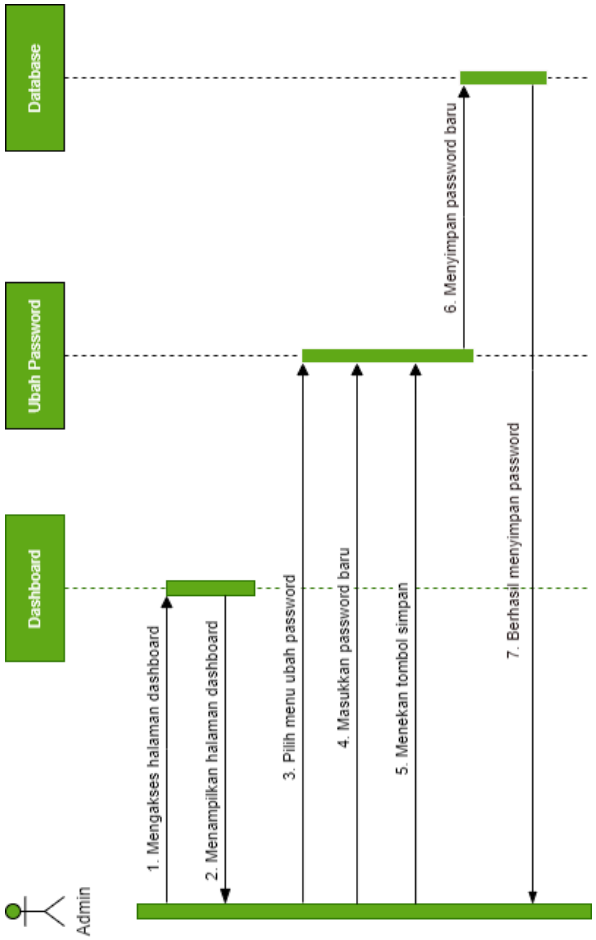

**Gambar III. 30** *Sequence* Lihat Laporan Hasil Prediksi

Pada Gambar III.30 menjelaskan proses *sequence* lihat laporan hasil prediksi yang dilakukan oleh Admin, Kaprodi, dan Dosen PA. Sistem akan mengakses data dan menampilkan hasil prediksi dihalaman laporan.



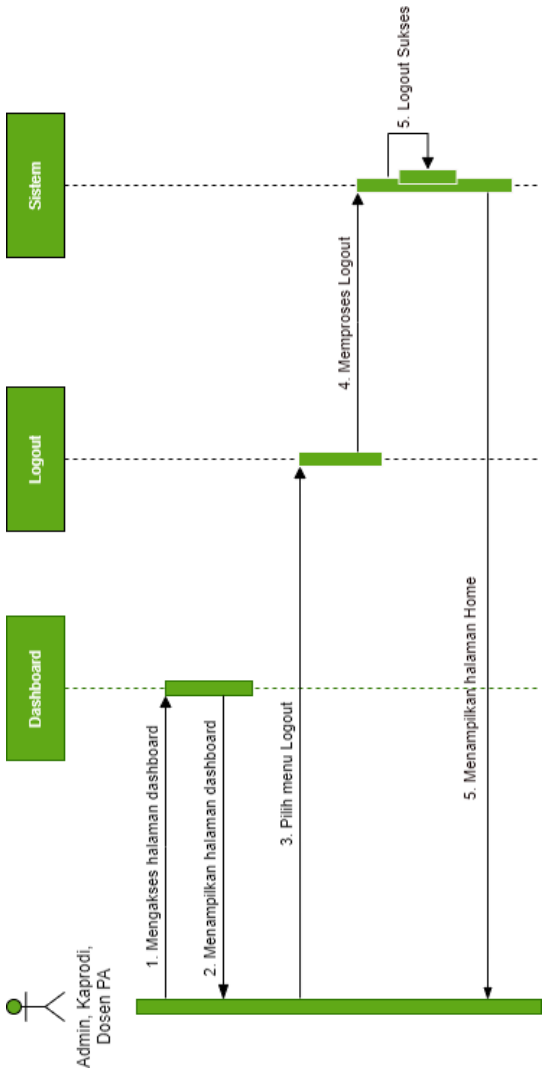

**Gambar III. 31** *Sequence* Cetak Laporan

Pada Gambar III.31 menjelaskan proses *sequence* cetak laporan yang dilakukan oleh admin. Sistem akan mengakses data dan admin akan memilih format data yang akan dicetak.




**Gambar III. 32** *Sequence* Ubah *Password*

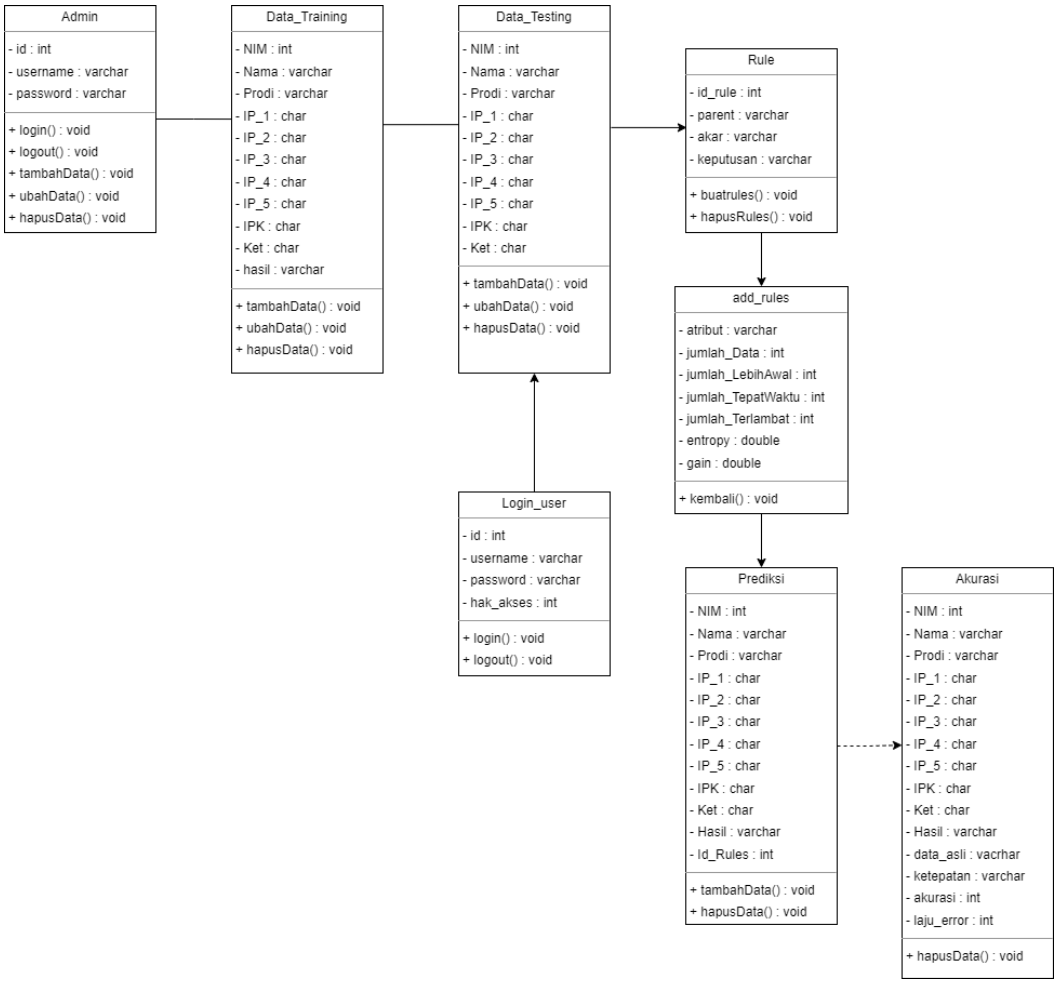
Pada Gambar III.32 menjelaskan proses *sequence* ubah *password* yang dilakukan oleh admin dengan mengisi form ubah *password. Password* yang telah diubah akan dikirim dan disimpan oleh sistem ke *database*.




**Gambar III. 33** *Sequence Logout*

Pada Gambar III.33 menjelaskan proses *sequence logout* yang dilakukan oleh Admin, Kaprodi, dan Dosen PA. Sistem akan memproses *logout* dan menampilkan halaman *home*.

### *III.5.7 Class Diagram*




**Gambar III. 34** *Class Diagram*

Pada Gambar III.34 menjelaskan *class diagram* pada sistem informasi prediksi kelulusan. Dapat dilihat bahwa admin dapat mengakses dan bertanggung jawab pada semua data dan proses prediksi. Sedangkan Kaprodi, dan Dosen PA, hanya dapat melihat data prodi, mencari data pada halaman mahasiswa, *rule*, target, dan melihat hasil prediksi.

## III.6 Populasi

Populasi merupakan variabel yang terkait dengan proses penelitian. Populasi dari penelitian ini meliputi data mahasiswa dari lima program studi yaitu Teknologi Informasi, Teknik Lingkungan, Arsitektur, Biologi, dan Kimia Angkatan 2020 yang diperoleh dari SIAKAD per Desember 2022-Maret 2023.

**Tabel III. 5** Populasi Angkatan 2020

| No | Populasi | Jumlah |
|---|---|---|
| 1. | Mahasiswa Kimia 2020 | 10 |
| 2. | Mahasiswa Biologi 2020 | 30 |
| 3. | Mahasiswa Arsitektur 2020 | 31 |
| 4. | Mahasiswa Teknik Lingkungan 2020 | 46 |
| 5. | Mahasiswa Teknologi Informasi 2020 | 77 |
| | **Jumlah Total Populasi** | 194 |

## III.7 Sampel

Sampel merupakan sebagian kecil dari jumlah dan karakteristik yang dimiliki oleh populasi dan dianggap mewakili seluruh populasi yang diperoleh menggunakan teknik tertentu. Berdasarkan jumlah populasi diatas, maka diambil sampel sebagai responden menggunakan rumus *slovin* :

Rumus :

$$n = \frac{N}{1+Ne^2} \text{.....Persamaan (3.5)}$$

Keterangan :

n = Sampel

N = Populasi

e = Batas toleransi kesalahan ( 5% )

1 = konstanta

a. Angkatan 2020

$$n = \frac{N}{1 + Ne^2}$$

$$n = \frac{194}{1 + (194 * (0{,}05^2))}$$

$$n = \frac{194}{1 + (194 * (0{,}0025))}$$

$$n = \frac{194}{1{,}485}$$

n = 130,63

n = 131 sampel

Dari perhitungan tersebut diketahui besar sampel penelitian ini adalah 131 sampel angkatan 2020.

## III.8 Teknik Pengambilan Sampel

Teknik pengambilan sampel yang ditentukan oleh peneliti yaitu dengan menggunakan *stratified random sampling*. *Stratified random sampling* merupakan sampel yang digunakan apabila ada perbedaan ciri antara setiap angkatan yang ada dan perbedaan tersebut mempengaruhi variabel. Setelah mendapatkan nilai sampel secara keseluruhan, maka langkah selanjutnya menentukan besaran sampel pada tiap kelas menggunakan rumus *propotional stratified random sampling* (Skripsi Cahyani Eka Putri) :

Rumus :

$$Spl = \frac{n}{N} * Js$$ ..... Persamaan (3.6)

Keterangan :

Spl : Jumlah sampel pada tiap-tiap sub populasi

N : Jumlah responden dalam populasi

n : Jumlah responden dalam sub populasi

Js : Jumlah sampel yang dibutuhkan (berdasarkan data sampel)

- Angkatan 2020
  1. Biologi

     $$Spl = \frac{30}{194}\ x\ 131 = 20{,}25 \text{ dibulatkan menjadi } 20$$

  2. Kimia

     $$Spl = \frac{10}{194}\ x\ 131 = 6{,}75 \text{ dibulatkan menjadi } 7$$

  3. Arsitektur

     $$Spl = \frac{31}{194}\ x\ 131 = 20{,}93 \text{ dibulatkan menjadi } 21$$

4. Teknik Lingkungan

$$Spl = \frac{46}{194} x\ 131 = 31{,}06 \text{ dibulatkan menjadi } 31$$

5. Teknologi Informasi

$$Spl = \frac{77}{194} x\ 131 = 51{,}99 \text{ dibulatkan menjadi } 52$$

**Tabel III. 6** Jumlah Sampel Mahasiswa Angkatan 2020

| No. | Program Studi | Jumlah |
|---|---|---|
| 1 | Kimia | 7 |
| 2 | Biologi | 20 |
| 3 | Arsitektur | 21 |
| 4 | Teknik Lingkungan | 31 |
| 5 | Teknologi Informasi | 52 |

## III.9 Jenis dan Sumber Data

Jenis data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh melalui *website* resmi Sistem Informasi Akademik (SIAKAD) dari lima program studi yang terdapat di Fakultas Sains dan Teknologi UIN Ar-Raniry. Data tersebut meliputi nama, nim, nama, IP semester 1-5, IPK, dan keterangan perkuliahan.

## III.10 Teknik Pengumpulan Data

Teknik pengumpulan data adalah cara yang digunakan dalam suatu penelitian untuk memperoleh data yang relevan. Teknik yang digunakan dapat mempengaruhi kualitas data yang dibutuhkan secara akurat dan tepat. Pada penelitian ini, peneliti menggunakan teknik pengumpulan data dokumentasi. Dokumentasi merupakan teknik yang digunakan untuk mencari data terkait hal-hal yang berkaitan dengan variabel penelitian. Dokumentasi dimanfaatkan peneliti sebagai penguat dan pendukung data yang didapatkan siakad prodi (Kurniawati, Annisa. 2022). Dokumen dalam penelitian ini berupa, IP semester 1-5, IPK, dan keterangan perkuliahan. Oleh karena itu, peneliti bekerjasama dengan admin

program studi Biologi, Kimia, Arsitektur, Teknik Lingkungan, dan Teknologi Informasi guna memperoleh data tersebut (Sulaeman, Ahmad, 2017).

## III.11 Teknik Pengolahan Data

Pengolahan data pada penelitian ini menggunakan *Microsoft Excel* 2019. Data yang diperoleh akan diurutkan sesuai dengan kebutuhan penelitian. Selain itu, akan dilakukan pengecekan berulang untuk memastikan data yang digunakan sudah sesuai dengan data sebenarnya.

## III.12 Teknik Analisis

Teknik analisis pada penelitian ini terbagi menjadi dua, yaitu teknik analisis data dan teknik analisis sistem.

### III.12.1 Teknik Analisis Data

Teknik analisis data yang digunakan pada penelitian ini adalah teknik kuantitatif deskriptif. Kuantitatif deskriptif merupakan teknik untuk menganalisis data dengan cara mendeskripsikan atau memberi gambaran terhadap objek yang diteliti melalui populasi atau sampel tertentu. Tujuannya untuk menggambarkan karakteristik suatu populasi dengan cara yang sistematis dan akurat (Nurcahyo, Arif Riska dkk. 2018). Menurut Sutinah (2013), penelitian kuantitatif memiliki 5 tahapan yang dilakukan secara konsisten :

1. Pemaparan latar belakang
2. Perumusan masalah penelitian
3. Mengemukakan tujuan penelitian
4. Mengemukakan teori yang digunakan dalam penelitian
5. Mengemukakan metodologi penelitian yang digunakan

Dibawah ini merupakan kondisi dan parameter dari variabel yang digunakan pada penelitian sebagai berikut :

1. IP (Indeks Prestasi)
   Variabel yang berisi total keseluruhan nilai IP dari semester 1 hingga semester 5 yang dimiliki oleh mahasiswa untuk diisi pada proses *input* program.

**Tabel III. 7** Predikat IP Program Studi S1

| IP | Predikat | Huruf |
|---|---|---|
| 3.50 – 4.00 | Pujian (*Cumlaude)* | A |
| 3.01 – 3.49 | Sangat Memuaskan | B |
| 2.76 – 3.00 | Memuaskan | C |
| 2.00 – 2.75 | Cukup | D |

2. IPK (Indeks Prestasi Kumulatif)

   Variabel yang berisi nilai IPK terakhir yang dimiliki oleh mahasiswa untuk diisi pada proses *input* program.

**Tabel III. 8** Predikat IPK Program Studi S1

| IP | Predikat | Huruf |
|---|---|---|
| 3.50 – 4.00 | Pujian (*Cumlaude)* | A |
| 3.01 – 3.49 | Sangat Memuaskan | B |
| 2.76 – 3.00 | Memuaskan | C |
| 2.00 – 2.75 | Cukup | D |

3. Keterangan Perkuliahan

   Variabel yang berisi data mahasiswa dengan pengelompokkan berdasarkan ketentuan yang dibuat pada program yaitu aktif, pernah mengajukan cuti, sedang cuti, dan non aktif.

**Tabel III. 9** Keterangan perkuliahan

| Ket | Huruf |
|---|---|
| Aktif | A |
| Pernah Cuti | B |
| Sedang Cuti | C |
| Non-Aktif | D |

4. Keputusan

   Variabel yang berfungsi untuk menentukan hasil keputusan. Dalam pengelompokkan data sudah ditentukan agar tidak terjadi kesalahan saat proses perhitungan program. Data keputusan terbagi menjadi 3 nilai yaitu "Lulus lebih awal", "Lulus tepat waktu", dan "Lulus terlambat".

III.12.2 Teknik Analisis Sistem

Analisis sistem informasi memiliki langkah-langkah dasar yang dapat diterapkan pada sistem yang dibangun yaitu sebagai berikut :

a. *Identify* : Mengidentifikasi masalah
b. *Understand* : Memahami kerja dari sistem
c. *Analyze* : Menganalisis sistem
d. *Report* : Membuat laporan hasil analisis

## III.13 Teknik Pengujian

Teknik pengujian yang digunakan pada penelitian ini terbagi menjadi 2, yaitu teknik pengujian data menggunakan *confusion matrix* dan teknik pengujian sistem menggunakan *Black Box testing.*

III.13.1 Teknik Pengujian Data

Pengujian data dilakukan untuk mengetahui akurasi dari data yang digunakan. Teknik pengujian data yang digunakan adalah *Confusion Matrix*. Terdapat tiga *performance matric* yaitu *accuracy, precision,* dan *recall* dengan ketentuan rumus sebagai berikut :

Rumus menghitung ***accuracy*** *confusion matrix* :

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \text{ x } 100\% \text{ ..... Persamaan (3.7)}$$

Rumus menghitung ***precision*** *confusion matrix* :

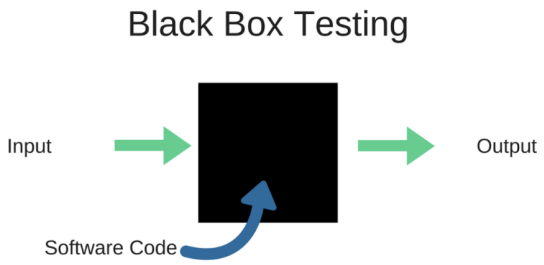
$$Precision = \frac{TP}{TP+FP} \text{ x } 100\% \text{ ..... Persamaan (3.8)}$$

Rumus menghitung ***recall*** *confusion matrix* :

$$Recall = \frac{TP}{TP+FN} \text{ x } 100\% \text{ ..... Persamaan (3.9)}$$

III.13.2 Teknik Pengujian Sistem

Pengujian sistem dilakukan untuk mengetahui cara kerja sistem, kelebihan, dan kekurangan dari sistem. Teknik pengujian yang digunakan adalah *Black box Testing*. *Black box testing* merupakan metode pengujian perangkat lunak yang berfokus pada keperluan *software* tanpa mengetahui struktur internal kode atau program ( TJ Sitinjak, Daniel Dido Jantce dkk, 2020).




**Gambar III. 35** Black box testing (Sumber : www.javadesde0.com)

Kesalahan-kesalahan yang sering ditemukan dengan menggunakan pengujian *black box testing* diantara lain :

*a.* Fungsi yang tidak benar atau hilang pada *software*
b. Kesalahan pada *interface* perangkat lunak
*c.* Kesalahan pada struktur data atau akses *database eksternal*
d. Kesalahan pada kinerja perangkat lunak
e. Kesalahan pada inisialisasi dan terminasi perangkat lunak.

# BAB IV

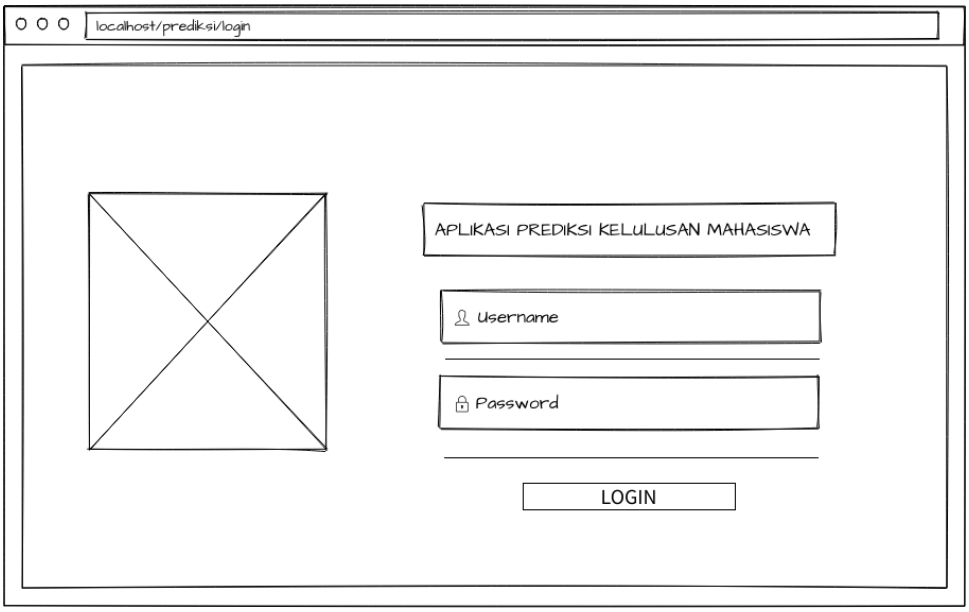
# HASIL DAN PEMBAHASAN

## IV.1 Perancangan Tampilan Antarmuka (*Interface*)

Rancangan antarmuka memberikan kemudahan kepada pengguna dalam melakukan interaksi dan komunikasi dengan sistem. Sistem memiliki tugas untuk menerima masukan dan menampilkan keluaran melalui antarmuka pada layar monitor. Pada penelitian ini, rancangan antarmuka dibagi menjadi tiga, yaitu rancangan menu, rancangan antarmuka masukan, dan rancangan antarmuka keluaran. (Azis,Nur. 2021).

### *1.* Halaman *Login*

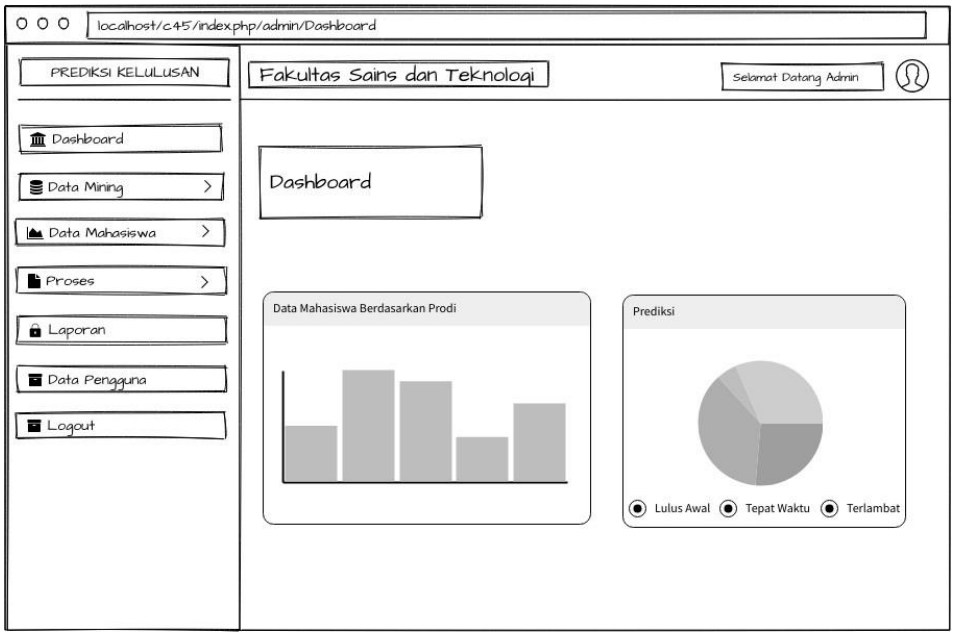
Pada aplikasi ini mempunyai tiga *user* yaitu admin, kaprodi, dan dosen penasehat akademik. Admin mempunyai hak akses memanagement data *training* dan perhitungan nilai *Entropy* dan *Gain* untuk mendapatkan *rule* prediksi. Sedangkan operator mempunyai hak akses *management* data target, tampil *rule* dan dan tampil laporan. Untuk dapat masuk ke halaman utama, *user* harus melakukan *login* terlebih dahulu dengan meng-*input username* dan *password*.



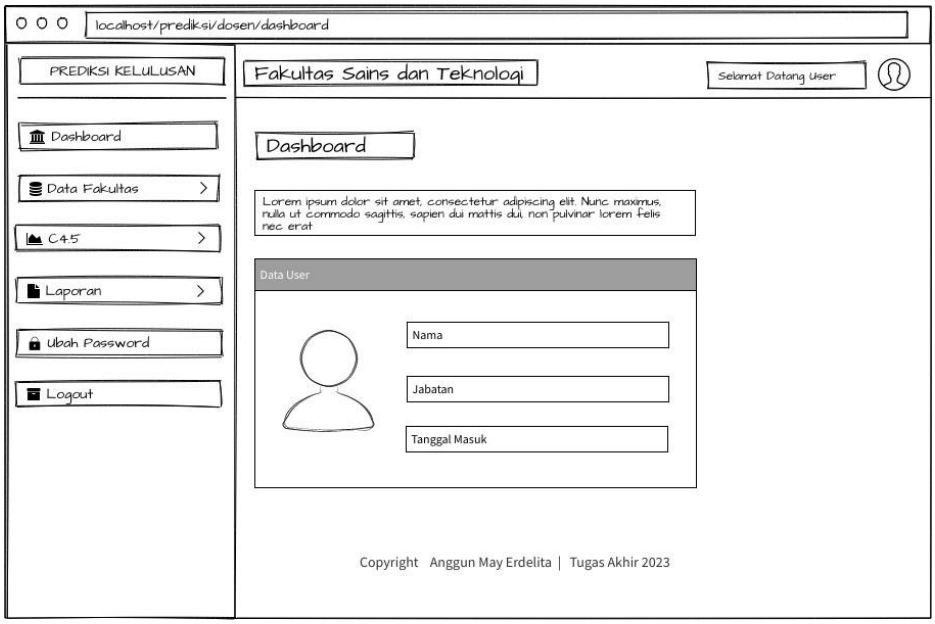

**Gambar IV. 1** Halaman *Login*

## 2. Halaman *Dashboard* (Utama)

Halaman *dashboard* atau halaman utama pada sistem informasi prediksi terdiri dari lima bagian yaitu data prodi, data mahasiswa, laporan, ubah *password*, dan *logout*. Halaman *dashboard* admin menampilkan diagram dari hasil prediksi kelulusan mahasiswa. Sedangkan halaman *dashboard user* menampilkan informasi singkat dari *user* yang mengakses sistem informasi tersebut. Berikut rancangan *interface* pada aplikasi :



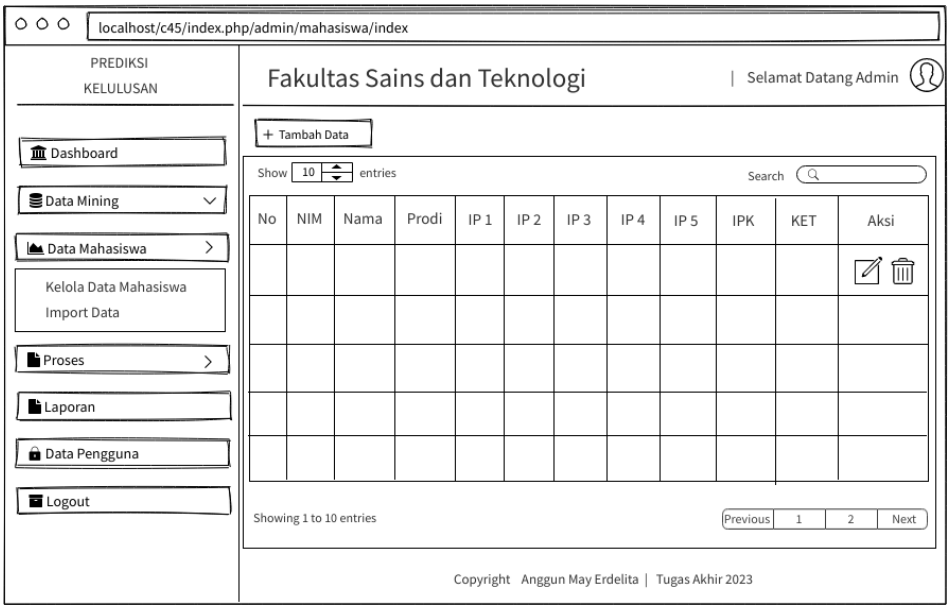

**Gambar IV. 2** Halaman *Dashboard* Admin




**Gambar IV. 3** Halaman *Dashboard User*

### *3.* Halaman Data *Testing*

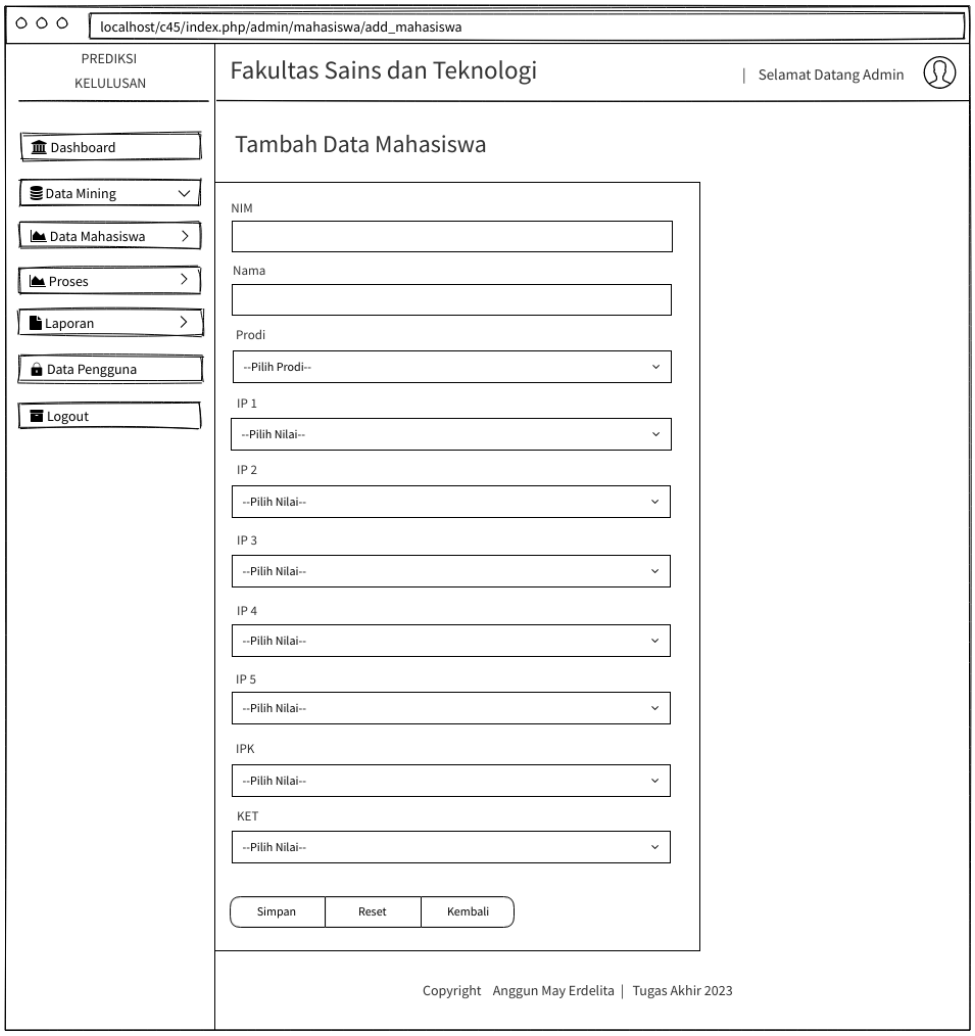
Halaman data mahasiswa merupakan data *testing* yang memiliki tiga fitur yaitu fitur tambah data mahasiswa, *update* data mahasiswa, dan hapus data. Berikut rancangan *interface* pada aplikasi :




**Gambar IV. 4** Halaman Data Mahasiswa

### *4.* Halaman Tambah Data *Testing*

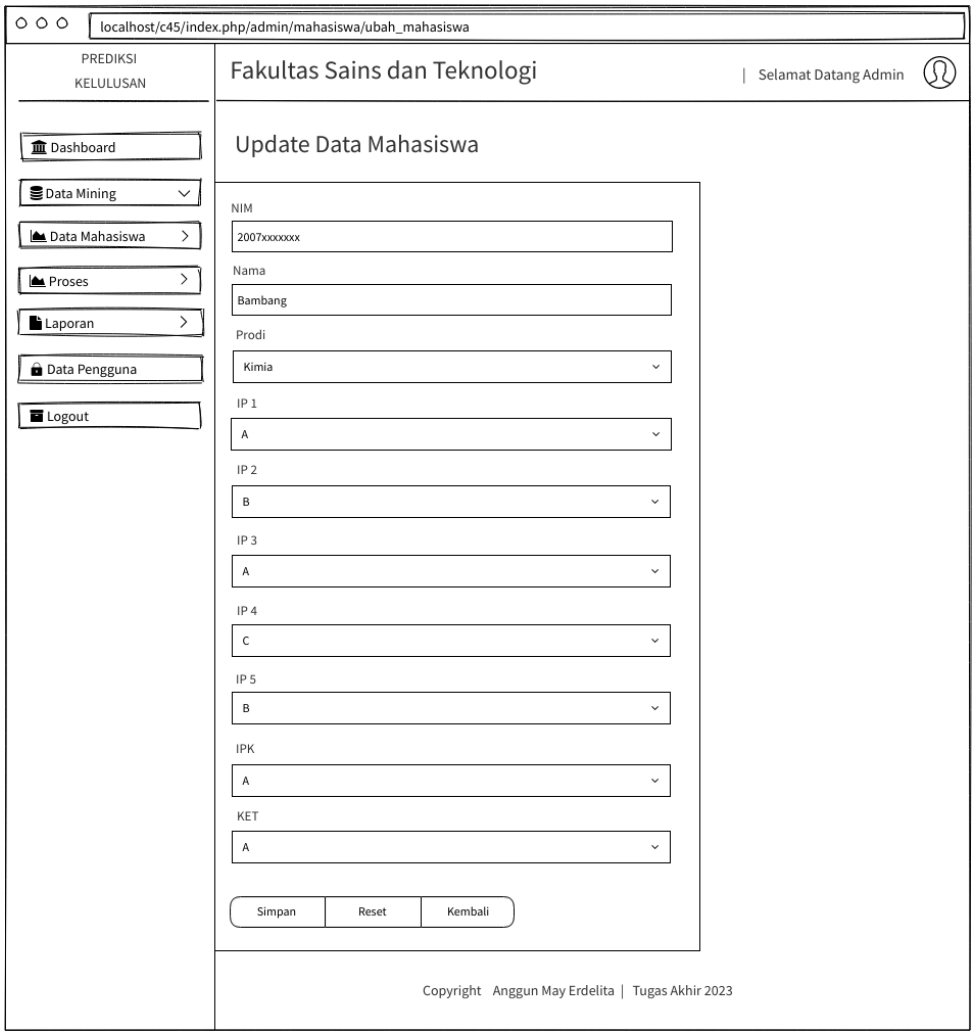
Halaman ini digunakan untuk memasukkan data mahasiswa baru dengan menginput nim, nama, prodi, IP 1, IP 2, IP 3, IP 4, IP 5, IPK, dan keterangan perkuliahan. Berikut rancangan *interface* pada aplikasi :




**Gambar IV. 5** Halaman Tambah Data Mahasiswa

### *5.* Halaman *Update* Data *Testing*

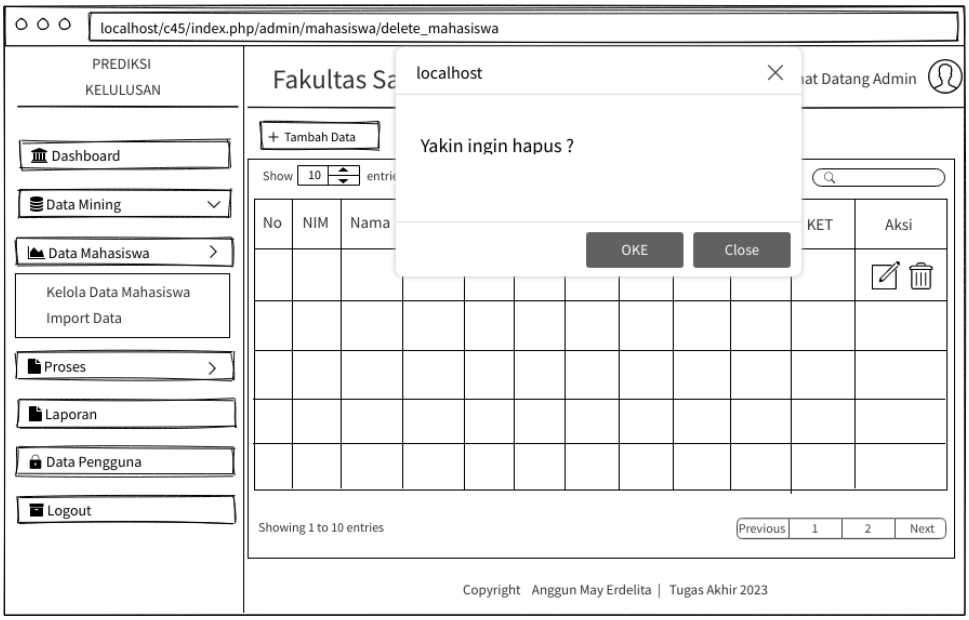
Halaman ini digunakan untuk mengubah data dengan menginput nim, nama, nim, nama, prodi, IP 1, IP 2, IP 3, IP 4, IP 5, IPK, dan keterangan perkuliahan. Berikut rancangan *interface* pada aplikasi :




**Gambar IV. 6** Halaman *Update* Data Mahasiswa

### *6.* Halaman Hapus Data *Testing*

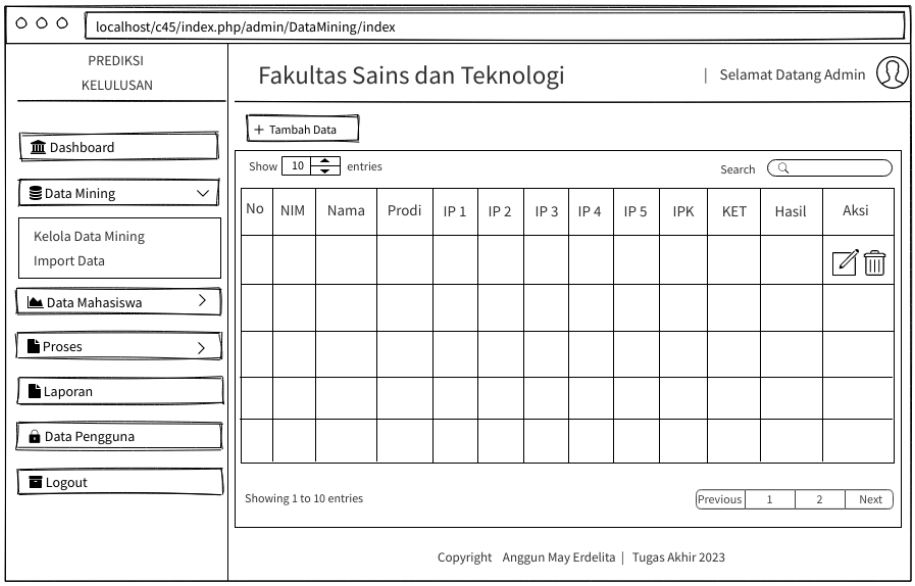
Halaman ini digunakan untuk menghapus data mahasiswa dengan memunculkan *pop-up* pemberitahuan . Berikut rancangan *interface* pada aplikasi :




**Gambar IV. 7** Halaman Hapus Data Mahasiswa

### *7.* Halaman Data *Training*

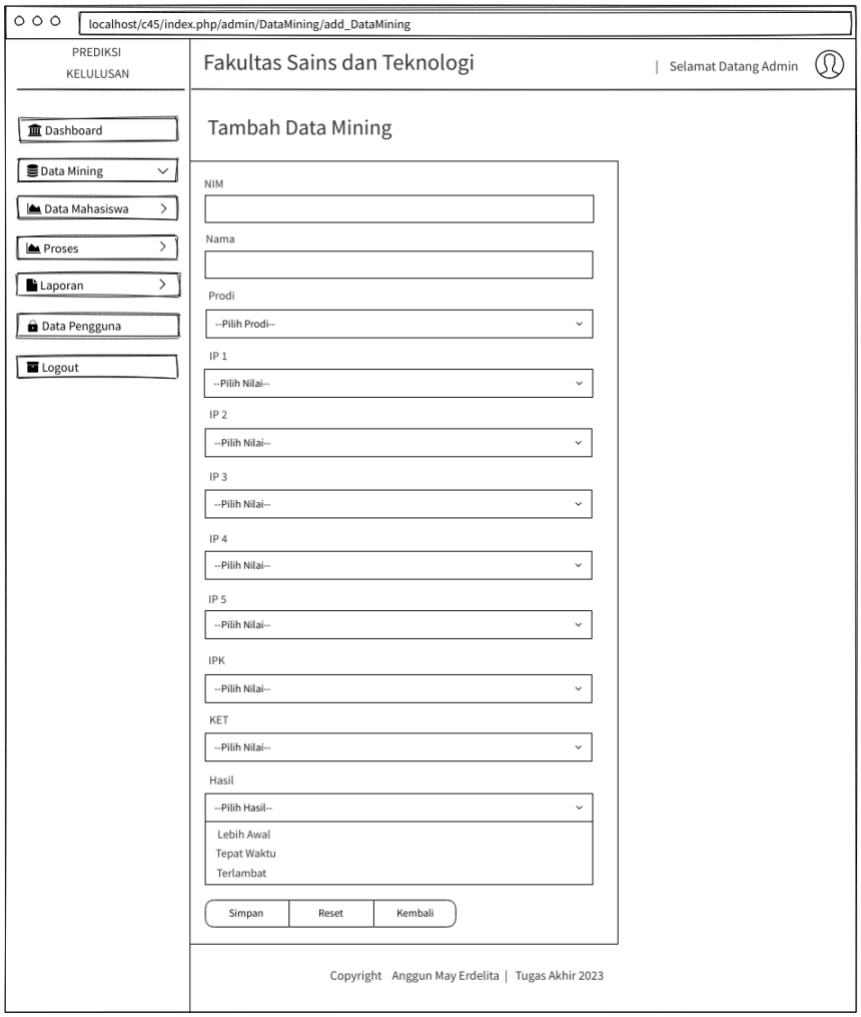
Halaman data mining merupakan halaman data training yang akan digunakan untuk prediksi kelulusan. Berikut rancangan *interface* pada aplikasi :




**Gambar IV. 8** Halaman Data Mining

### *8.* Halaman Tambah Data *Training*

Halaman ini digunakan untuk memasukkan data *training* baru dengan meng-*input* nim, nama, prodi, IP 1, IP 2, IP 3, IP 4, IP 5, IPK, keterangan perkuliahan, dan hasil prediksi awal. Berikut rancangan *interface* pada aplikasi:




**Gambar IV. 9** Halaman Tambah Data *Training*

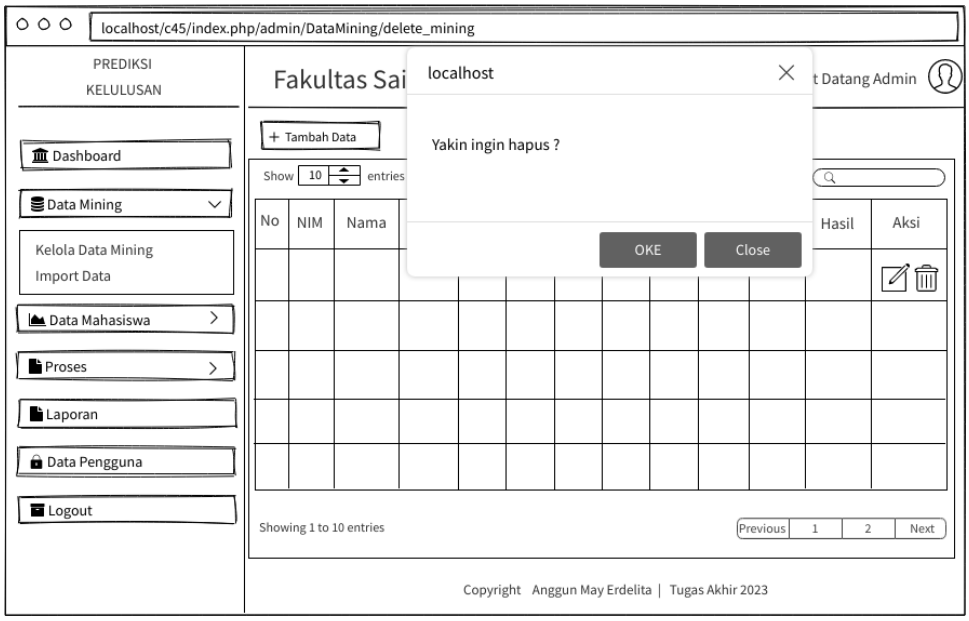
### *9.* Halaman *Update* Data *Training*

Halaman ini digunakan untuk mengubah data dengan meng-*input* nim, nama, nim, nama, prodi, IP 1, IP 2, IP 3, IP 4, IP 5, IPK, dan keterangan perkuliahan. Berikut rancangan *interface* pada aplikasi :

**Gambar IV. 10** Halaman Ubah Data *Training*

### *10.* Halaman Hapus Data *Training*

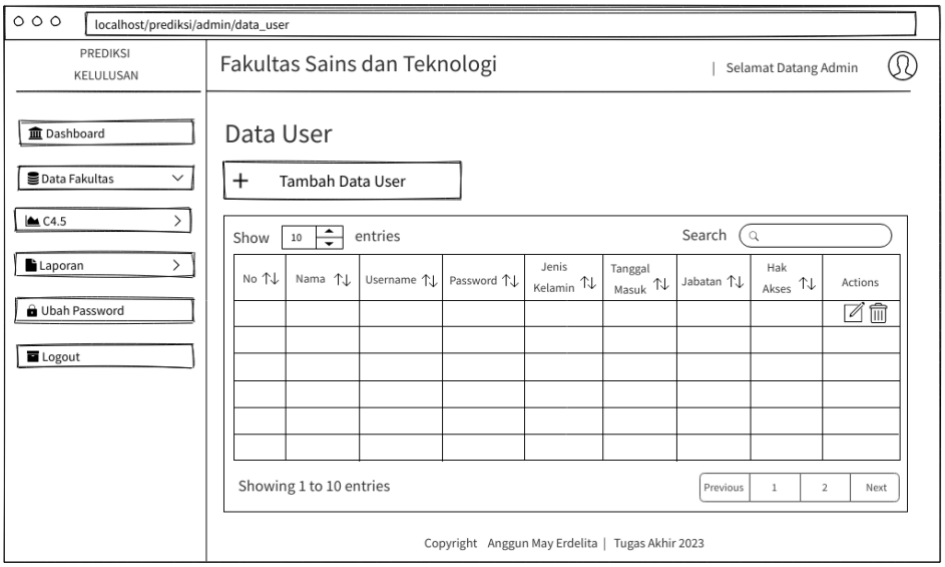
Halaman ini digunakan untuk menghapus data training dengan memunculkan *pop-up* pemberitahuan. Berikut rancangan *interface* pada aplikasi:




**Gambar IV. 11** Halaman Hapus Data *Traning*

### *11.* Halaman Data *User*

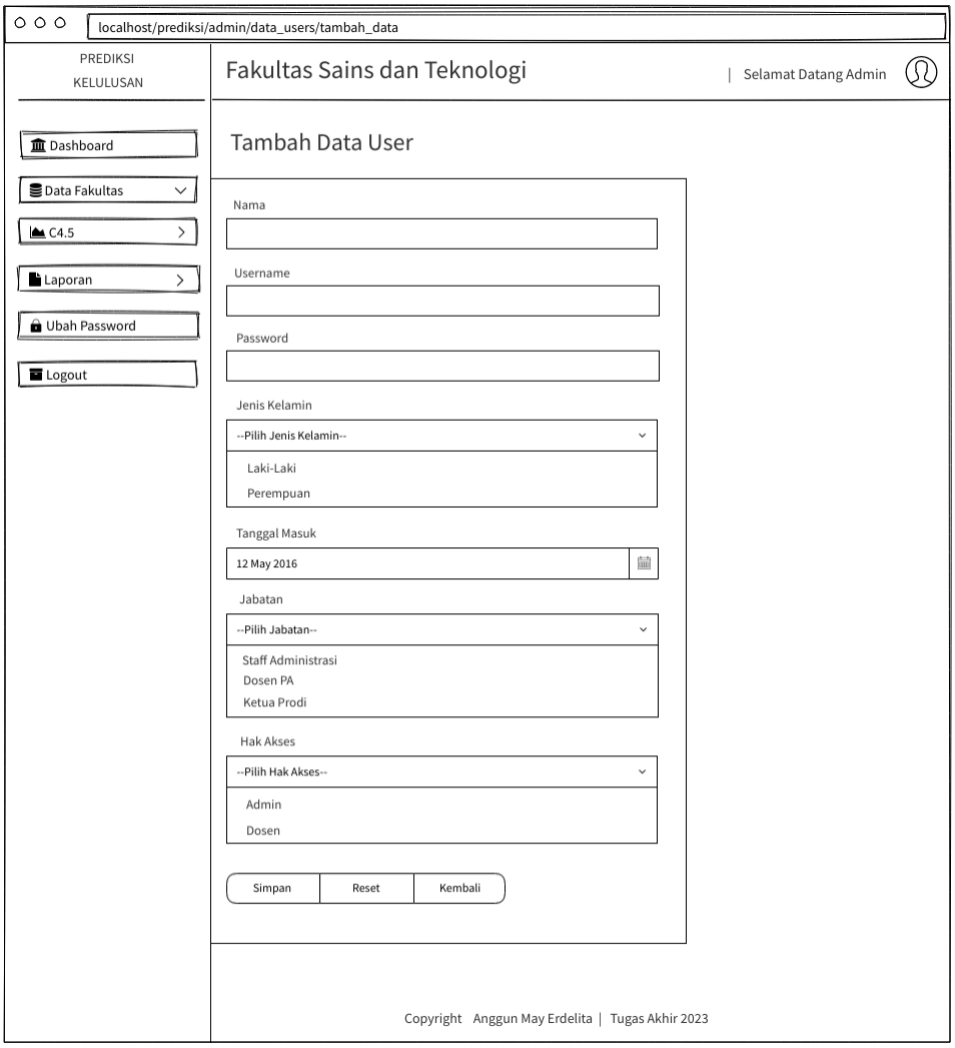
Halaman ini menampilkan data user yang melakukan aktivitas *login* ke dalam sistem informasi prediksi kelulusan yang terdiri dari nama, *username*, *password*, jenis kelamin, tanggal masuk, jabatan, hak akses, fitur tambah data, cari data, *update* data, dan hapus data. Beirkut rancangan *interface* pada aplikasi :




**Gambar IV. 12** Halaman Data *User*

### *12.* Halaman Tambah Data *User*

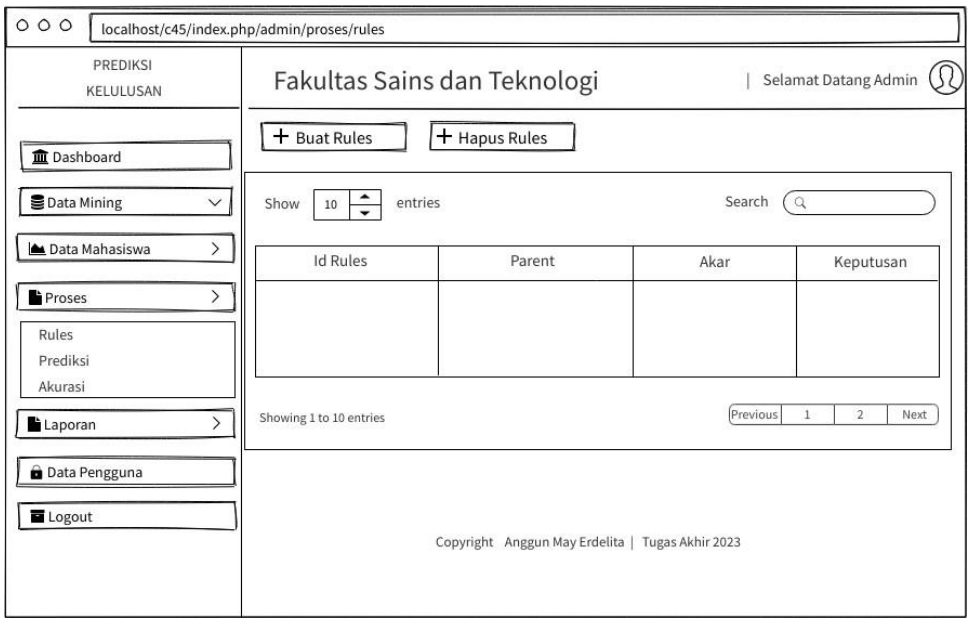
Halaman ini digunakan untuk memasukkan data *user* baru yang terhubung ke halaman *login* dengan meng-*input* nama, *username, password*, jenis kelamin, tanggal masuk, jabatan, status, dan hak akses. Berikut rancangan *interface* pada aplikasi :



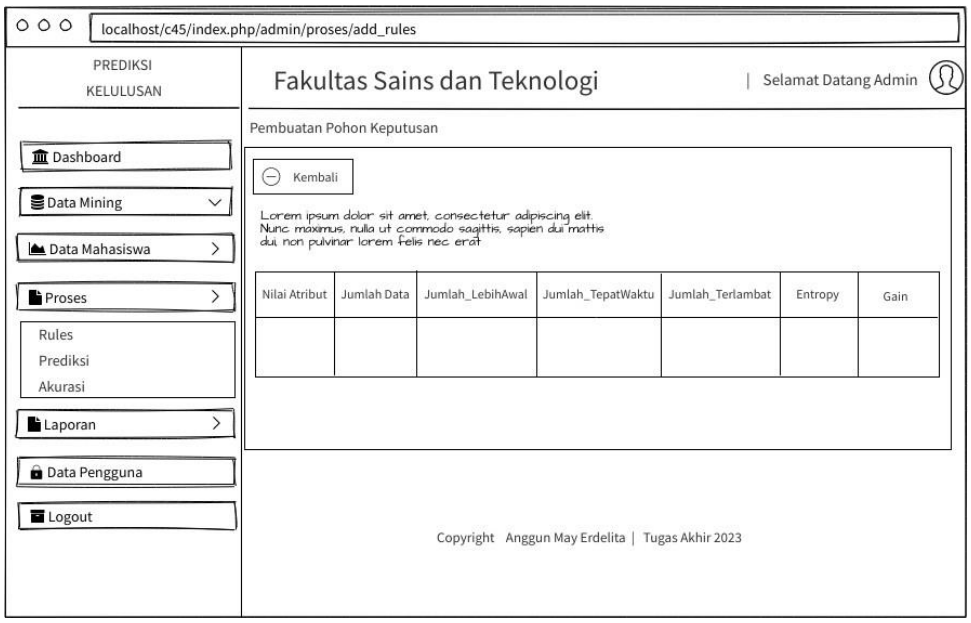

**Gambar IV. 13** Halaman Admin Tambah Data *User*

### *13.* Halaman *Rules*

Halaman rules merupakan halaman untuk pembentukan pohon keputusan, perhitungan *entropy* dan *gain*. Nilai *entropy* dan *gain* dari hasil perhitungan tiap atribut pada data *training* menggunakan algoritma C4.5.



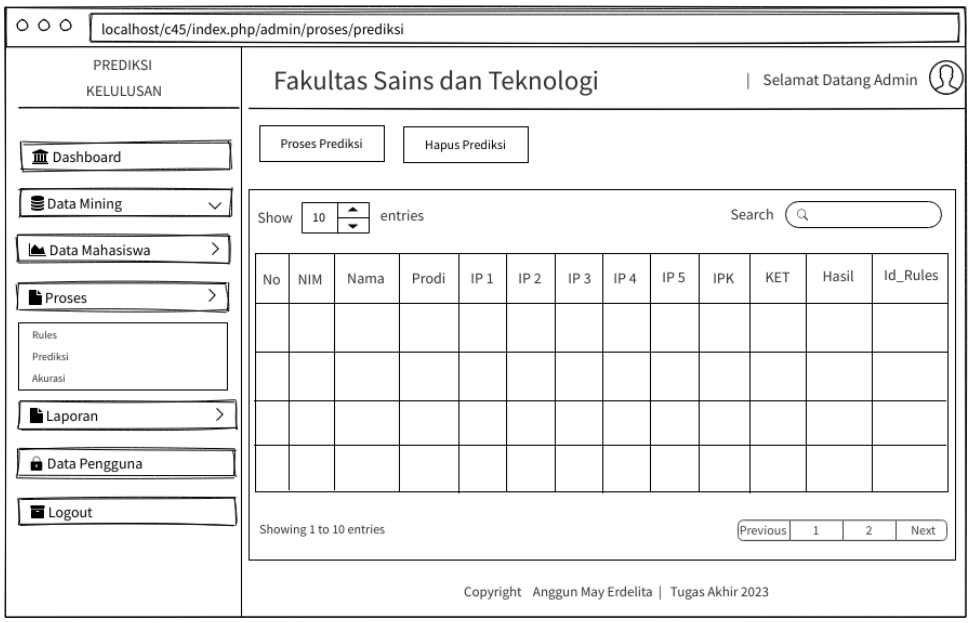

**Gambar IV. 14** Halaman *Rules*




**Gambar IV. 15** Halaman *Entropy* dan *Gain*

## 14. Halaman Prediksi

Halaman prediksi merupakan halaman untuk melakukan proses prediksi kelulusan dengan menampilkan nim, nama, prodi, IP semester 1-5, IPK, Keterangan perkuliahan, hasil prediksi, dan Id *rules* yang telah terbentuk. Berikut rancangan *interface* pada aplikasi :




**Gambar IV. 16** Halaman Proses Prediksi Kelulusan

## 15. Halaman *Import* Data

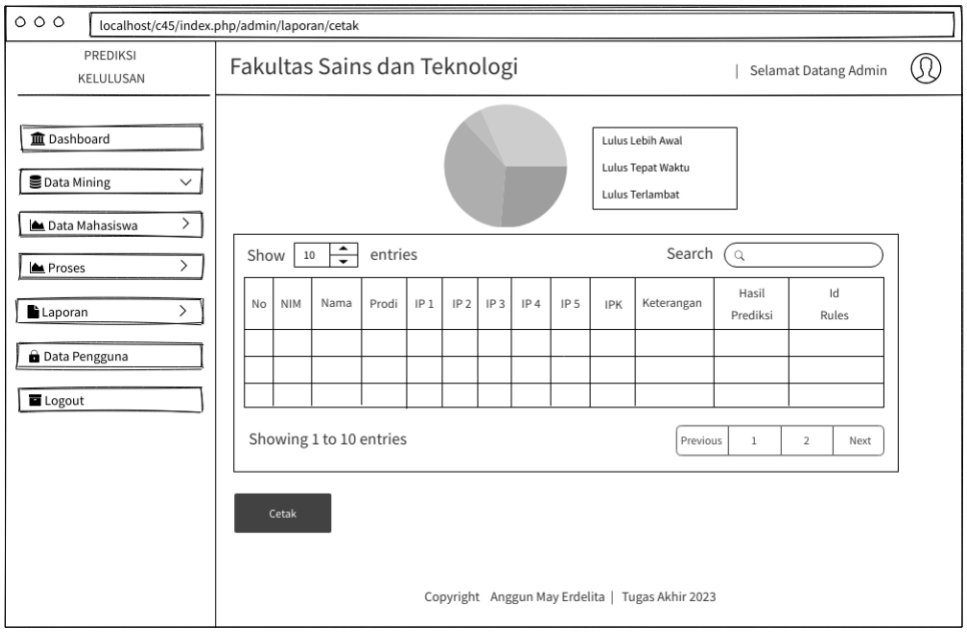
Halaman ini digunakan untuk meng-*import* data dari *Microsoft Excel* ke dalam sistem yang nantinya data tersebut akan ditampilkan dalam tabel data *training*. Berikut rancangan *interface* pada aplikasi :

**Gambar IV. 17** Halaman *Import* Data *Training*

**Gambar IV. 18** Halaman *Import* Data *Testing*

## 16. Halaman Laporan

Halaman laporan berisi tentang hasil prediksi kelulusan yang juga menampilkan hasil berupa diagram. Selain itu, terdapat fitur cetak laporan hasil prediksi. Berikut rancangan *interface* pada aplikasi :




**Gambar IV. 19** Halaman Cetak Laporan

### 17. Halaman Ubah Password

Halaman ini digunakan untuk mengganti *password* dengan cara menginput *password* yang baru dan mengulangi *password* yang telah diinput. Setelah itu, klik untuk menyimpan perubahan *password* tersebut.

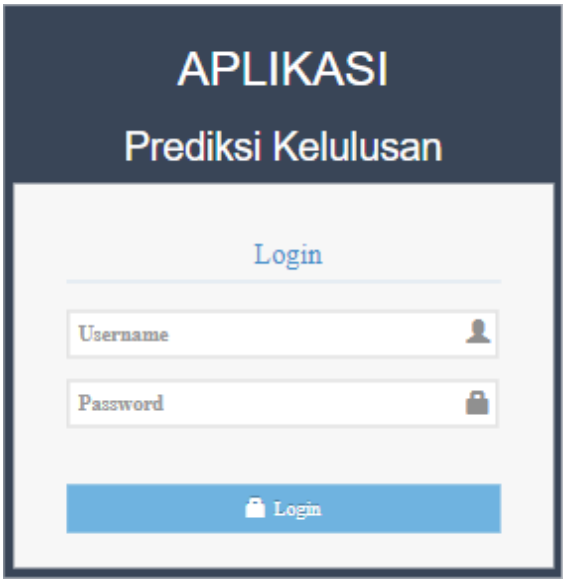
**Gambar IV. 20** Halaman Ubah *Password*

## IV.2 Implementasi Sistem

Sistem informasi prediksi kelulusan yang dihasilkan dari penelitian ini dibangun dengan menggunakan bahasa pemrograman *PHP7.3*, *framework Codeigniter* 3.1.0, dan database *MySQL*. Tahapan implementasi untuk memprediksi kelulusan mahasiswa lebih awal, tepat waktu, dan terlambat meliputi perhitungan data *training*, proses data mining C4.5, pembentukan *rules*, proses prediksi, menghitung akurasi, dan menampilkannya ke dalam bentuk diagram.

*1.* Halaman *Login*

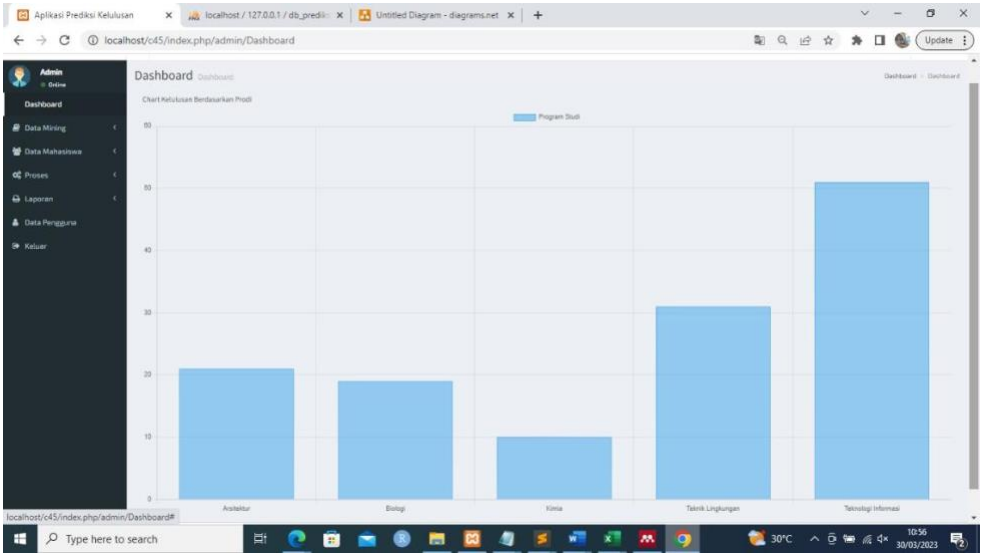
Sistem informasi prediksi kelulusan ini terdapat tiga *user* yaitu admin, ketua prodi, dan dosen penasehat akademik yang hak akses berbeda. Berikut tampilan halaman *login* pada program.




**Gambar IV. 21** Halaman *Login*

2. Halaman *Dashboard*

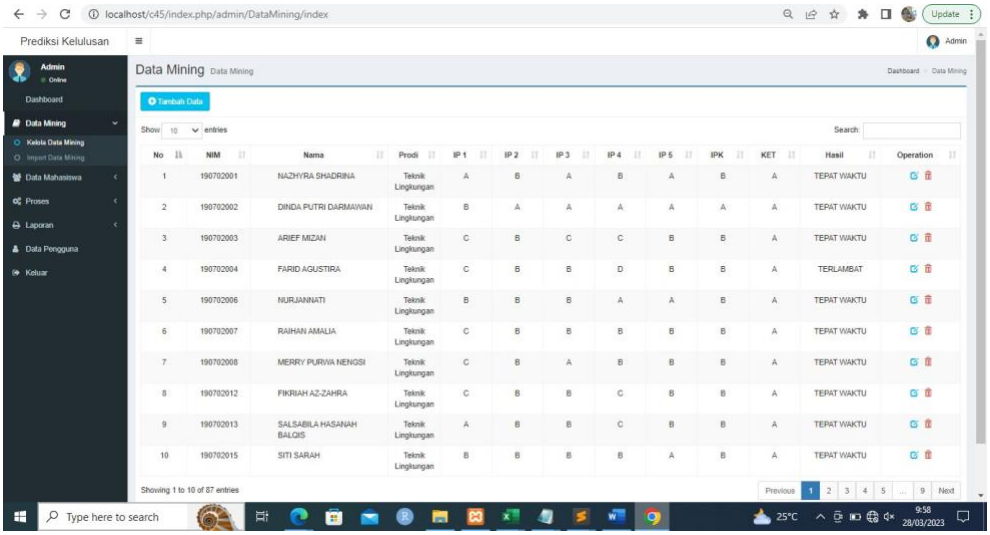
Tampilan *dashboard* pada Gambar 4.22 merupakan halaman *dashboard* admin. Admin memiliki hak akses yang dapat melakukan semua aktivitas termasuk perhitungan *entropy* dan *gain*.




**Gambar IV. 22** Halaman *Dashboard*

3. Halaman Data *Training*

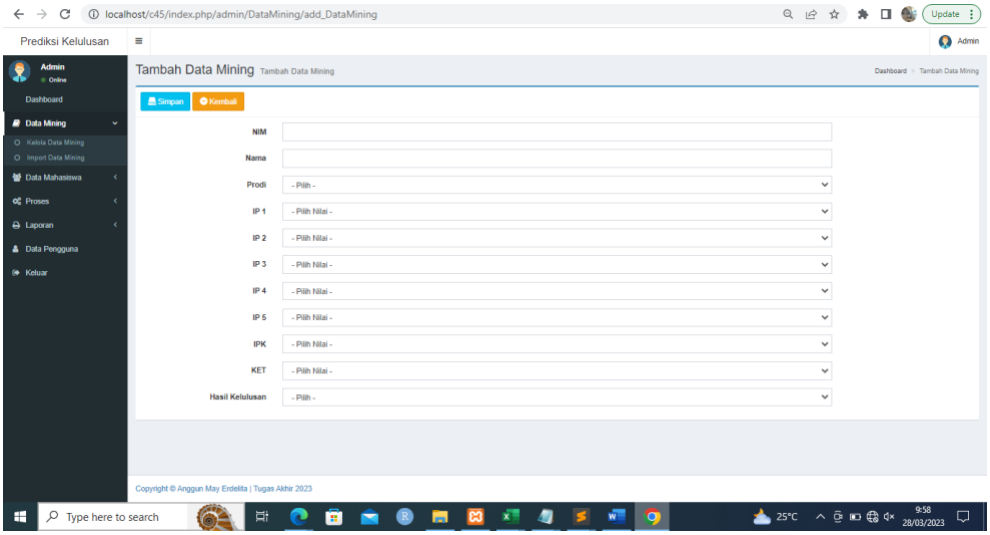
Halaman data mining dibawah ini merupakan halaman data sampel *training* yang terdiri dari beberapa atribut yaitu nim, nama, prodi, IP 1, IP 2, IP 3, IP 4, IP 5, IPK dan keterangan kelulusan (hasil prediksi). Pada halaman ini atribut dikelompokkan dalam bentuk tabel yang nantinya akan dihitung nilai *entropy* dan *gain*-nya.



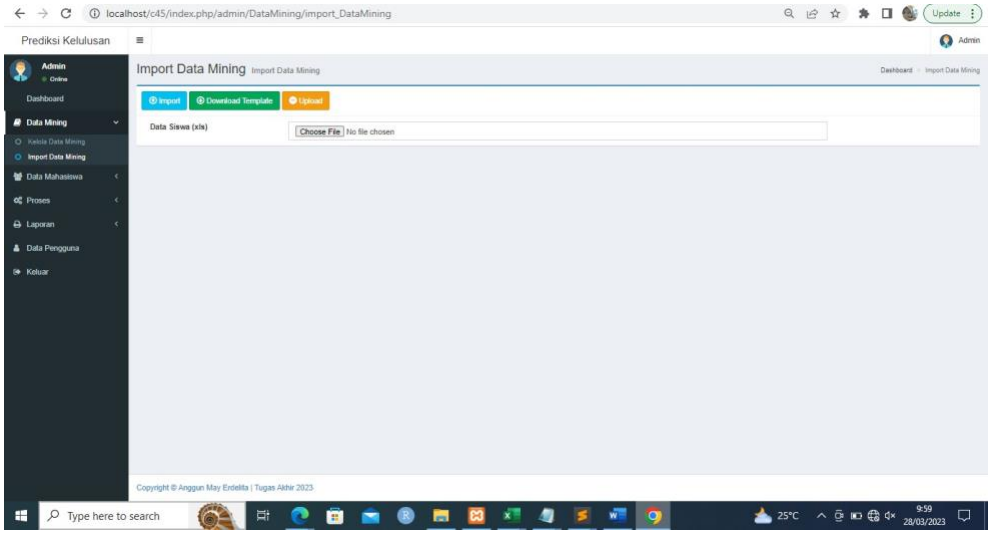

**Gambar IV. 23** Data *Training*

*4.* Halaman *Input* dan *Import* Data *Training*

Halaman *input* data *training* dapat digunakan untuk menambahkan data ke dalam tabel secara individu, sedangkan halaman *import* digunakan untuk memasukkan banyak data ke dalam sistem dalam satu waktu dengan format *excel*. Berikut tampilannya seperti gambar dibawah ini :



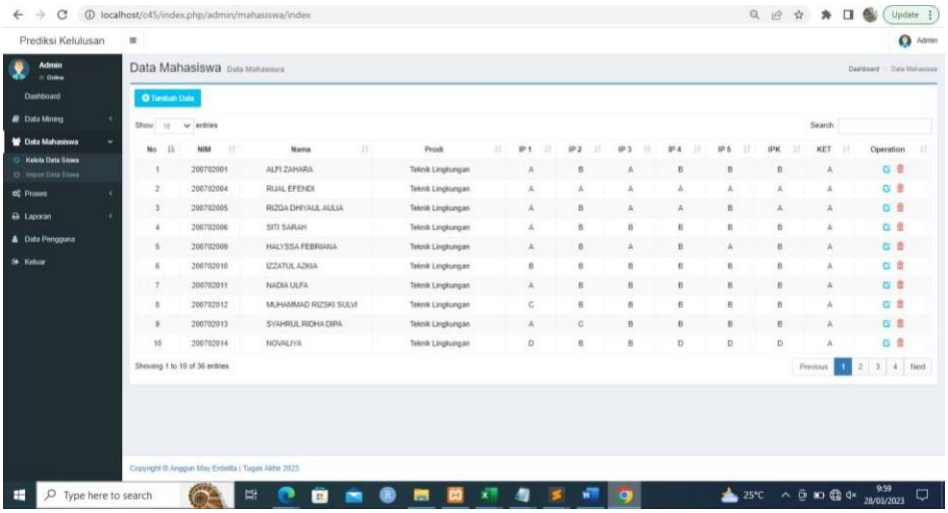

**Gambar IV. 24** Input Data *Training*



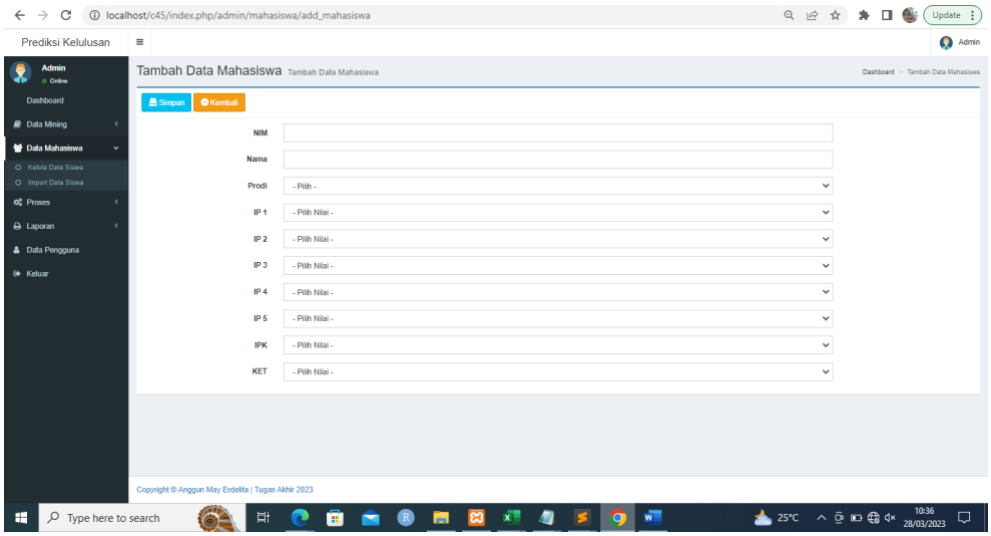

**Gambar IV. 25** Import Data *Training*

5. Halaman Data *Testing*

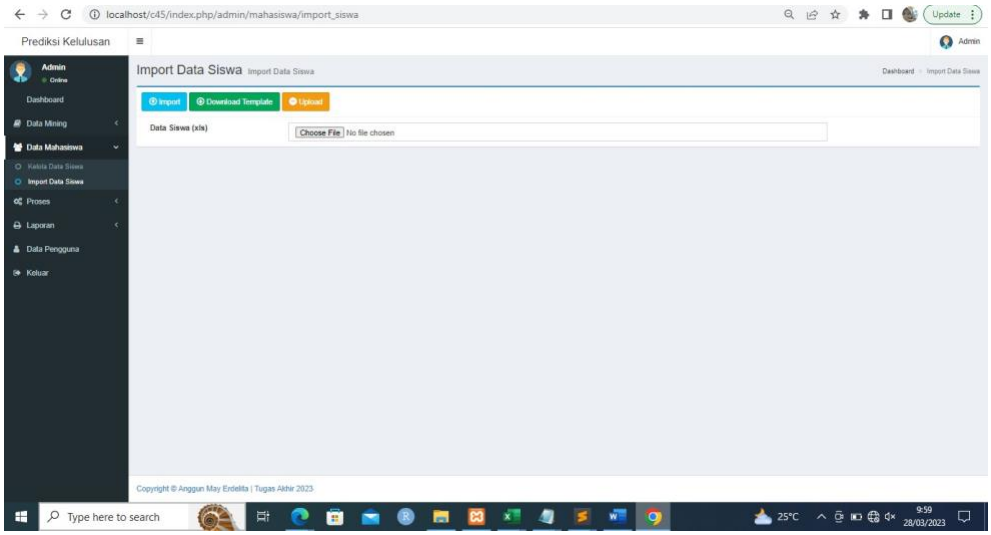
Halaman data *testing* merupakan halaman data uji yang akan diprediksi dengan memasukkan nilai sesuai atributnya. Data *testing* dapat ditambahkan baik secara individu melalui form tambah data atau melakukan *import* data dari *excel*. Berikut tampilan gambarnya dibawah ini :




**Gambar IV. 26** Data *Testing*



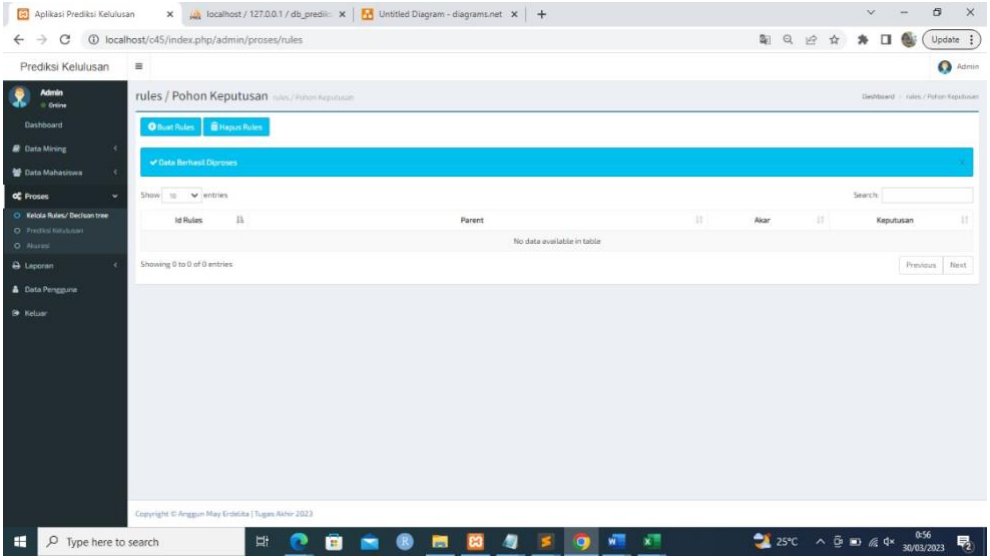

**Gambar IV. 27** Input Data *Testing*




**Gambar IV. 28** Import Data *Testing*

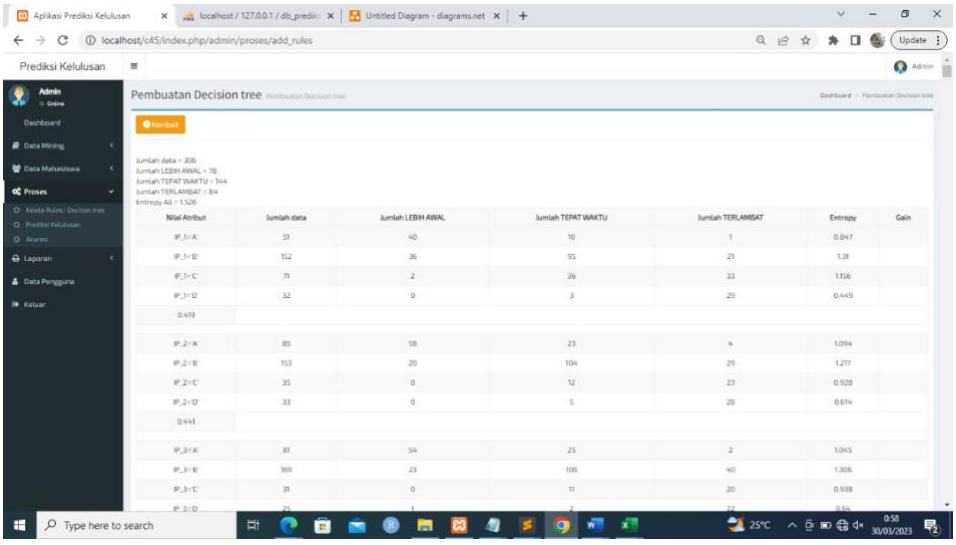
*6.* Halaman Pembentukan *Rules*

Halaman *rule* merupakan halaman aturan yang berisi kondisi dari data dalam pennetuan keputusan. Aturan tersebut di dapatkan berdasarkan perhitungan *entropy* dan *gain* dari data *training*.




**Gambar IV. 29** Pohon Keputusan

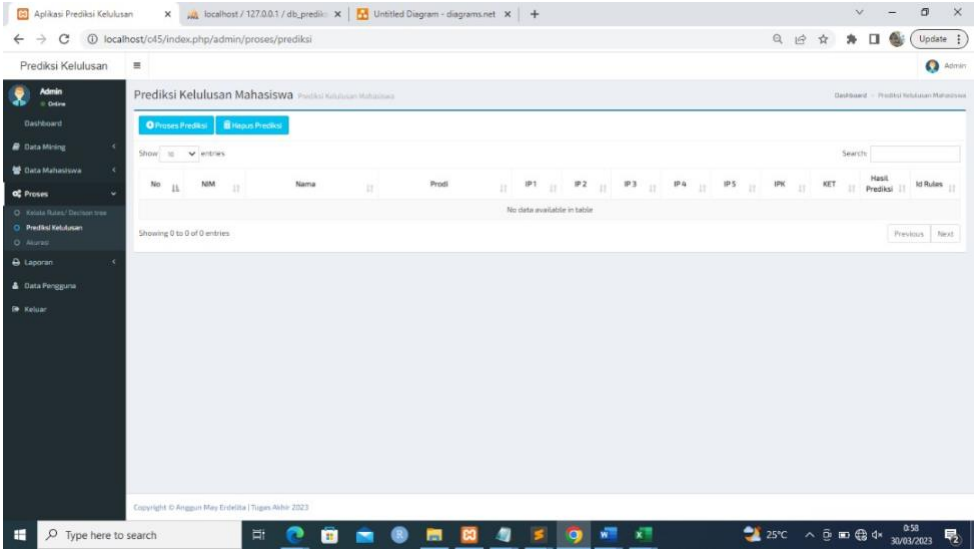
Proses pembentukan pohon keputusan memiliki perulangan perhitungan berdasarkan perolehan data *training*. Perulangan pertama menghasilkan nilai *gain* tertinggi yang digunakan untuk mengklasifikasikan kasus menjadi aturan lulus lebih awal (X), tepat waktu (Y), dan terlambat (N). kemudian akan dilakukan perulangan selanjutnya, tetapi atribut dengan *gain* tertinggi yang diperoleh dari perhitungan sebelumnya tidak dimasukkan lagi.




**Gambar IV. 30** Perhitungan *Entropy & Gain*

7. Halaman Proses Prediksi

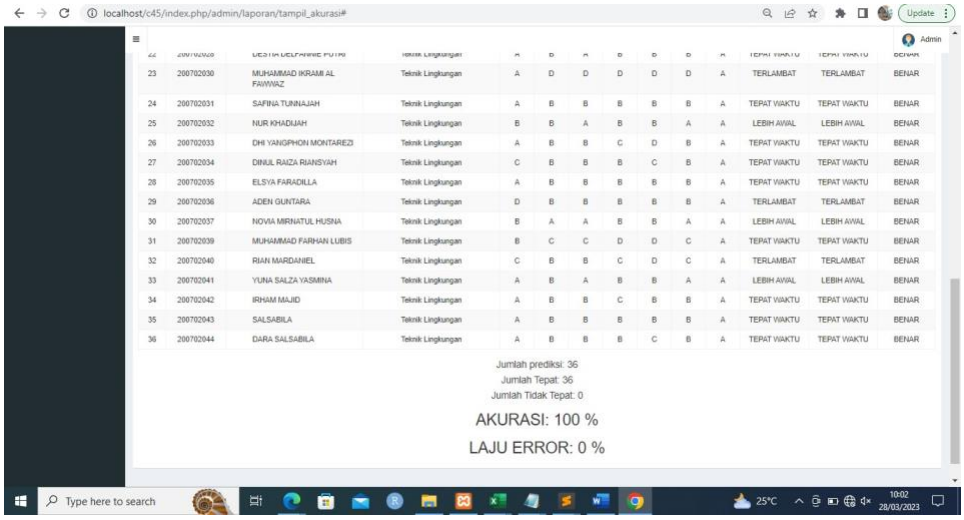
Halaman dibawah ini merupkan proses prediksi dari data *testing* yang digunakan dengan menampilkan hasil keputusan berdasarkan aturan *id rules* yang di dapat.




**Gambar IV. 31** Proses Prediksi

8. Halaman Akurasi Prediksi

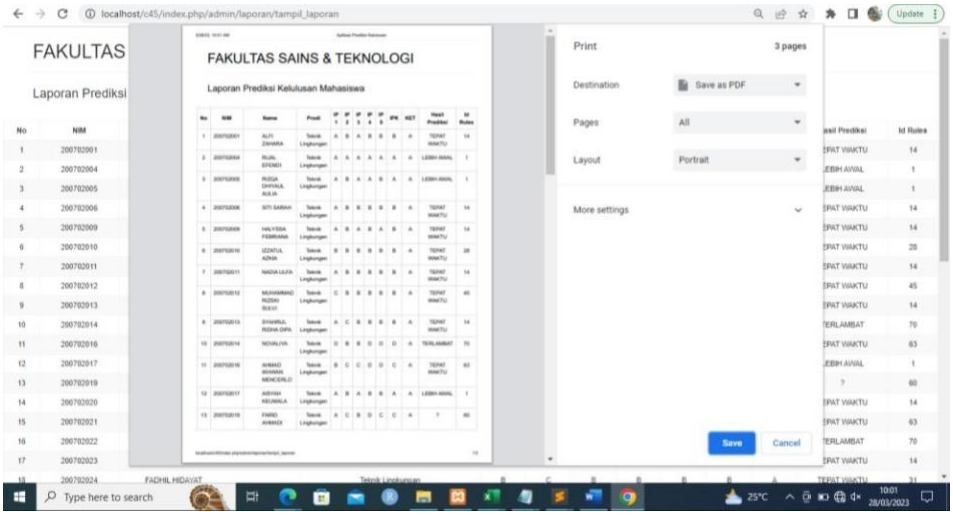
Halaman akurasi pada gambar dibawah ini menampilkan hasil akurasi dari proses prediksi yang telah selesai dilakukan dengan membandingkan nilai dari data *testing* dengan data *training*.



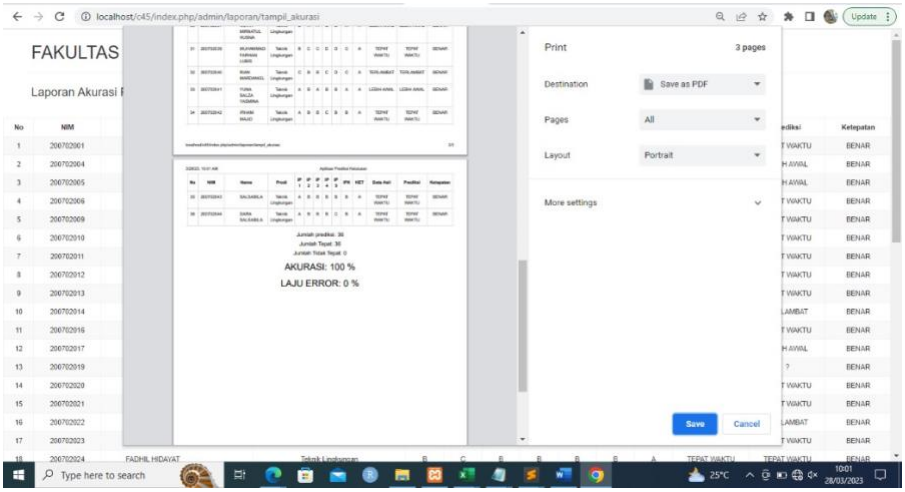

**Gambar IV. 32** Proses Akurasi Data

9. Halaman Cetak Laporan

Halaman cetak laporan terbagi menjadi dua. Pertama, laporan dapat dicetak langsung tanpa menapilkan hasil akurasi prediksi. Kedua, laporan dapat dicetak dengan menampilkan hasil akurasi diakhir tabel.



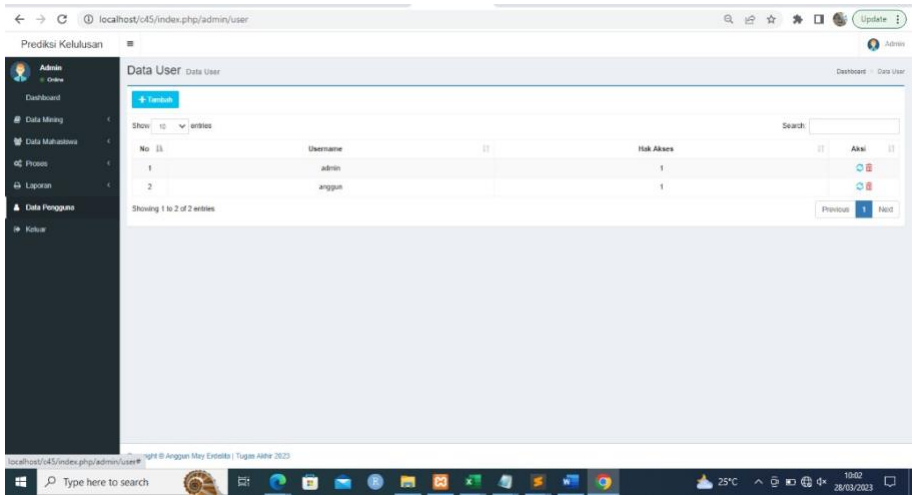

**Gambar IV. 33** Cetak Laporan Prediksi



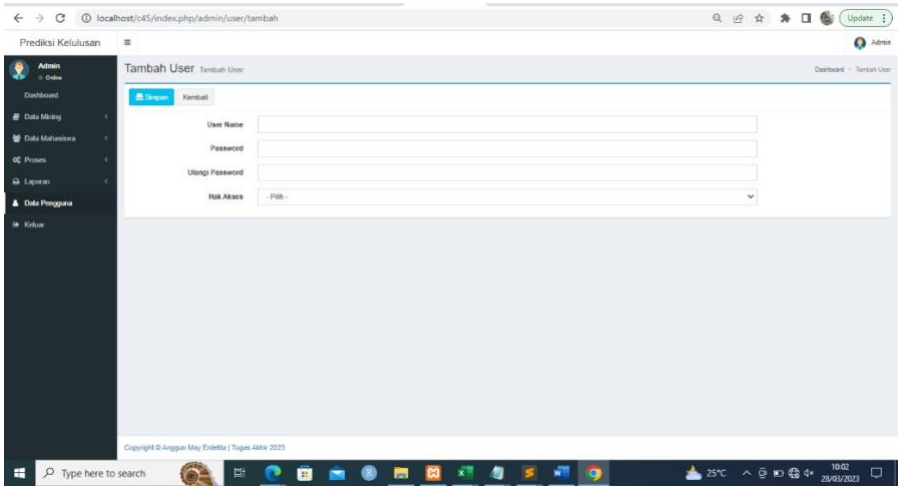

**Gambar IV. 34** Cetak Laporan Akurasi Prediksi

10. Halaman Data Pengguna

Halaman data pengguna merupakan halaman *user* yang terdiri dari *username* dan hak akses. Data pengguna hanya dapat ditambahkan oleh admin dengan menginput data *user* pada *form* inputan. Berikut tampilan gambar dibawah ini :




**Gambar IV. 35** Halaman *User*




**Gambar IV. 36** Input Data *User*

## IV.3 Dataset

Data merupakan fakta atau data mentah yang berhubungan dengan konteks permasalahan. Dataset pada penelitian ini berupa kumpulan data mahasiswa meliputi IP semester I-V, IPK terakhir, dan keterangan perkuliahan. Data yang dikumpulkan diambil dari *website* SIAKAD UIN Ar-Raniry. Proses pemilihan data menggunakan *Microsoft Excel*.

Untuk mendapatkan hasil pengujian deteksi yang baik dan memiliki keakuratan tinggi dibutuhkan data *training* yang banyak untuk dilakukan proses *training* agar sistem dapat mengenali dan mempelajari pola data tersebut. Dataset dibagi menjadi data *training* 80% dan data *testing* 20%. Pembagian data dapat dilihat pada tabel

**Tabel 4. 1** Data *Training*

| Label | Jumlah Data |
|---|---|
| A | 771 |
| B | 928 |
| C | 252 |
| D | 190 |
| Total | 2.141 (80%) |

**Tabel 4. 2** Data *Testing*

| Label | Jumlah Data |
|---|---|
| A | 267 |
| B | 383 |
| C | 136 |
| D | 133 |
| Total | 919 (20%) |

### *IV.4 Confusion Matrix*

*Confusion matrix* digunakan untuk menghitung keakuratan pada sebuah objek deteksi. Berikut perhitungan *confusion matrix* pada data yang telah diuji menggunakan system.

**Tabel 4. 3** *Confusion Matrix*

| | | Actual | | |
|---|---|---|---|---|
| | | Lebih Awal | Tepat Waktu | Terlambat |
| Predicted | Lebih Awal | 22 | 3 | 0 |
| | Tepat Waktu | 0 | 49 | 4 |
| | Terlambat | 0 | 3 | 4 |
| Total | | 22 | 55 | 8 |

Berikut ini proses perhitungan *accuracy*, *precision*, dan *recall* :

Persamaan *Accuracy* = $\frac{Jumlah\ objek\ terdeteksi\ benar}{Jumlah\ keseluruhan\ objek\ yang\ terdeteksi}$ x 100%

= $\frac{75}{85}$ x 100%

= 88%

Persamaan *Precision* = $Precision = \frac{TP}{(TP + FP)}$ x 100%

= $\frac{75}{75+3}$ x 100 %

= $\frac{75}{78}$ x 100 %

= 96 %

Persamaan *Recall* = $Recall = \frac{TP}{(TP + FN)}$ x 100%

= $\frac{75}{75+4}$ x 100 %

= $\frac{75}{79}$ x 100 %

$= 94\ \%$

Persamaan *F1 Score* $= F1\ Score = \frac{1}{2}\ (\frac{1}{Presis} + \frac{1}{Recall})$ x 100%

$= \frac{1}{2}\ (\frac{1}{88} + \frac{1}{94})$ x 100%

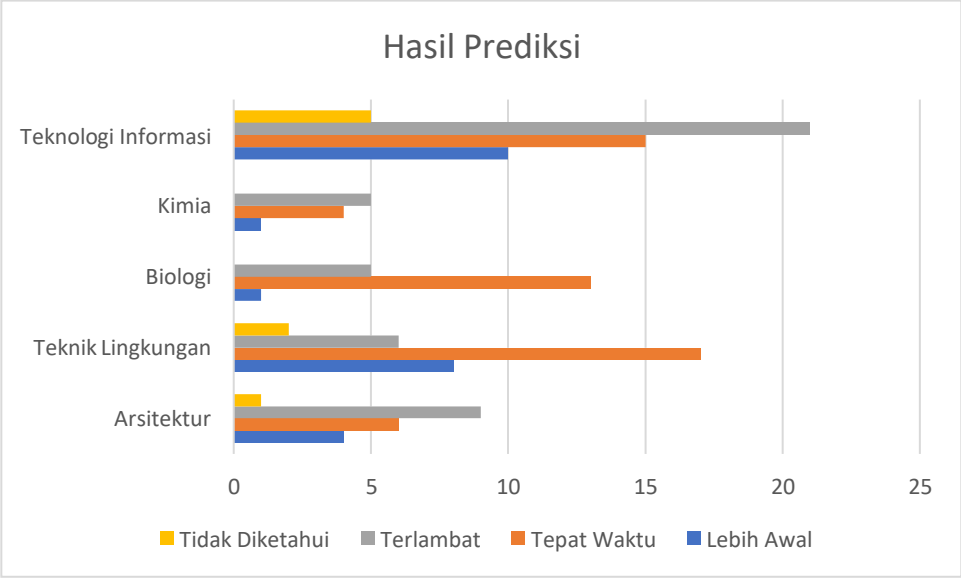
$= \frac{1}{2}\ (\frac{2}{182})$ x 100%

$= \frac{1}{91}$ x 100%

$= \frac{1}{91}\ \%$

***Tabel 4. 4*** Hasil Perhitungan *Accuracy, Precision, dan Recall* dan *F1 Score* Pada Data *Testing*

| ***Confusion Matrix*** | **Data Testing** |
|---|---|
| ***Accrucay*** | 88% |
| ***Precision*** | 96% |
| ***Recall*** | 94% |
| ***F1 Score*** | 91% |




**Gambar IV. 37** Hasil Prediksi

Berdasarkan gambar 4.37 diatas diperoleh hasil bahwa kelulusan lebih awal dan tepat waktu lebih mendominasi. Banyaknya data *testing* dari setiap program studi sangat mempengaruhi hasil prediksi kelulusan. Seperti pada Program Studi Teknologi Informasi dengan jumlah data sebanyak 51 memiliki hasil prediksi terlambat paling besar diantara program studi lainnya. Akan tetapi, terdapat 8 data yang tidak dapat diprediksi oleh sistem. Hal ini dapat disebabkan oleh pembentukan id *rule* yang masih belum terbentuk secara sempurna. Selain itu, hasil pengujian sistem menggunakan *black box testing* menghasilkan sebuah kesimpulan bahwa masih terdapat fitur yang belum berfungsi yaitu fitur *import* data dari *excel*, dan akses halaman *user*.

Dengan demikian, penerapan algoritma prediksi kelulusan ke dalam sistem informasi *website* menggunakan bahasa pemrograman *PHP* dan *framework Codeigniter* dengan keluaran tiga ouput yaitu lulus lebih awal, tepat waktu, dan terlambat memiliki hasil *accuracy* 88%, *precision* 96%, dan *recall* 94%. Atribut yang paling berpengaruh dalam proses prediksi adalah IPK. Hasil penelitian ini memiliki kemiripan dengan penelitian terdahulu yang diteliti oleh (Mulia, Isnan & Muanas, 2021) dalam memprediksi kelulusan mahasiswa di Universitas Peradaban dengan hasil *accuracy* sebesar 88%, *precision* 91,79%, dan *recall* 95%, tetapi memiliki perbedaan dengan hasil keluaran yang dihasilkan.

# BAB V

# KESIMPULAN DAN SARAN

## V.1 Kesimpulan

Dari hasil penelitian yang telah dilakukan dan dipaparkan diatas, penulis mengambil beberapa kesimpulan pada penelitian ini, yaitu :

1. Dataset yang digunakan pada penelitian sebanyak 437 data. Data *training* memiliki jumlah 306 data, dan data *testing* berjumlah 131 data. Dari data tersebut diperoleh tingkat *accuracy* sebesar 88%, *precision* 96%, dan *recall* 94%.
2. Berdasarkan hasil prediksi dapat dilihat bahwa prodi yang memiliki kelulusan lebih awal dan terlambat terbanyak adalah Teknologi Informasi. Sedangkan prodi yang memiliki kelulusan tepat waktu terbanyak dari prodi Teknik Lingkungan. Hal ini dapat dipengaruhi karena jumlah data *testing* dari setiap program studi yang berbeda. Selain itu, terdapat 5 data lainnya yang tidak dapat diprediksi. *Id rule* yang terbentuk pada sistem sebanyak 169 *rule* yang menjadi pohon keputusan dalam memprediksi.

## V.2 Saran

Dalam penelitian ini, terdapat beberapa saran dari penulis, diantaranya sebagai berikut :

1. Penulis mengharapkan saran perbaikan dan pengembangan dalam memperbaiki fitur *import* dan halaman akses untuk *user*, serta pembentukan id *rule* agar semua data dapat terbaca dengan baik.
2. Menambahkan dataset agar data yang diprediksi memiliki hasil yang lebih akurat.
3. Menerapkan algoritma lain untuk dapat melihat perbandingan hasil akurasi prediksi.

# DAFTAR PUSTAKA


Anwar F.F, Jaya.A.I, & Abu, M. (2022). Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Decision Tree dengan Penerapan Algoritma C4.5. *Jurnal Ilmiah Matematika Dan Terapan*, *19*(1), 19–28. https://doi.org/10.22487/2540766x.2022.v19.i1.15880

Azis, N. (2021). Perbandingan dan Prediksi Kelulusan Mahasiswa Dengan Weka. In *Angewandte Chemie International Edition, 6(11), 951–952.*

Daniel Dido, Maman, J. S. (2020). Analisa Dan Perancangan Sistem Informasi Administrasi Kursus Bahasa Inggris Pada Intensive English Course Di Ciledug Tangerang. *JURNAL IPSIKOM Vol. 8 No.1, Juni 2020 ISSN : 2338-4093, E-ISSN : 2686-6382*, *8*(1).

Gholamy, A., Kreinovich, V., & Kosheleva, O. (2018). Why 70/30 or 80/20 Relation Between Training and Testing Sets : A Pedagogical Explanation. *Departmental Technical Reports (CS)*, *1209*, 1–6.

Hermawanti, S. N., Asriyanik, & Sunarto, A. A. (2019). Implementasi Algoritma C4.5 untuk Prediksi Kelulusan Tepat Waktu ( Studi Kasus : Program Studi Teknik Informatika ). *Jurnal Ilmiah SANTIKA*, *9*(1), 853–864. http://jurnalummi.agungprasetyo.net/index.php/santika/article/download/552/253

Kurniawati, A. (2022). Pengaruh Adversity Quotient dan Optimisme Hasil IPK Terhadap Studi Tepat Waktu Mahasiswa PAI Tingkat Akhir UIN Malang. *Material Safety Data Sheet*, *33*(1), 1–12. http://www2.warwick.ac.uk/fac/sci/whri/research/mushroomresearch/mushroomquality/fungienvironment%0Ahttps://us.vwr.com/assetsvc/asset/en_US/id/16490607/contents%0Ahttp://www.hse.gov.uk/pubns/indg373hp.pdf

Mashlahah, S. (2013). *Prediksi Kelulusan Mahasiswa Menggunakan Metode Decision Tree Dengan Penerapan Algoritma C4.5*.

Mulia, I., & Muanas, M. (2021). Model Prediksi Kelulusan Mahasiswa Menggunakan Decision Tree C4.5 dan Software Weka. *JAS-PT (Jurnal*

*Analisis Sistem Pendidikan Tinggi Indonesia)*, *5*(1), 71. https://doi.org/10.36339/jaspt.v5i1.417

Mustika, Ardilla, Y., Manuhutu, A., Ahmad, N., Hasbi, I., Guntoro, Manuhutu, M. A., Ridwan, M., Hozairi, Wardhani, A. K., Alim, S., Romli, I., Religia, Y., Octafian, D. T., Sufandi, U. U., & Ernawati, I. (2021). *Data Mining Dan Aplikasinya* (Vol. 1). www.penerbitwidina.com

Nathan, A. J., & Scobell, A. (2012). Model Algoritma K-nearest Neighbor untuk memprediksi kelulusan mahasiswa. *Foreign Affairs*, *91*(5), 1–9.

Nurcahyo, A. R. (2018). *Makalah Desain Penelitian Kuantitatif Deskriptif. 17070855416*, 16 hlm. https://www.academia.edu/36311890/MAKALAH_DESAIN_PENELITIAN_KUANTITATIF_DESKRIPTIF

Pardosi, W., Hutagalung, D. M., Damanik, B., & Ginting, R. U. (2020). *Rancang Bangun Sistem Ujian Semester Computer Based Test (Cbt) Dan Implementasi Di Kab. Deliserdang (Studi Kasus Di Smk Negeri 1 Kutalimbaru)*. *6*(2), 89–93. https://garuda.kemdikbud.go.id/documents/detail/2018568

Putri, R. P. S., & Waspada, I. (2018). Penerapan Algoritma C4.5 pada Aplikasi Prediksi Kelulusan Mahasiswa Prodi Informatika. *Khazanah Informatika: Jurnal Ilmu Komputer Dan Informatika*, *4*(1), 1. https://doi.org/10.23917/khif.v4i1.5975

Rahman, A. F. A., Sorikhi, & Wartulas, S. (2020). Prediksi Kelulusan Mahasiswa Menggunakan Algoritma C4.5 (Studi Kasus di Universitas Peradaban). *IJIR (Indonesian Journal of International Relation)*, *1*(2), 70–77.

Riyadi, F. A. (2020). *Implementasi Metode naive Bayes Untuk Prediksi Kelulusan Mahasiswa Tepat Waktu Prodi Informatika (Studi Kasus : Universitas Teknologi Yogyakarta)*. 1–9. http://eprints.uty.ac.id/4863/

Rohman, A., & Rufiyanto, A. (2019). Implementasi Data Mining Dengan Algoritma Decision Tree C4 . 5 Untuk Prediksi Kelulusan Mahasiswa Di Universitas Pandaran. *Proceeding SINTAK 2019*, 134–139.

Sulaeman, A. (2017). *Pengaruh Keaktifan Mahasiswa Dalam Mengikuti Organisasi Himpunan Mahasiswa Pendidikan Administrasi Perkantoran (HIMA ADP) dan Disiplin Belajar Terhadap Prestasi Belajar Mahasiswa Program Studi Pendidikan Administrasi Perkantoran Fakultas Ekonomi UNY*. 1–14.

Suweleh Abdurraghib Segaf, Dyah Susilowati, H. (2020). Aplikasi Penentuan Penerima Beasiswa Menggunakan Algoritma C4.5. *Jurnal BITe*, *2*(1), 12–21. https://doi.org/10.30812/bite.v2i1.798

Widhianto, F. P., Supianto, A. A., & Setiawan, N. Y. (2019). Aplikasi Data Mining Untuk Memprediksi Kelulusan Mahasiswa Menggunakan Algoritme C4.5. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, *3*(6), 5549–5555. http://files/179/Widhianto et al. - Aplikasi Data Mining Untuk Memprediksi Kelulusan M.pdf

Yuningsih, L., Setiawan, I. R., & Sunarto, A. A. (2020). Rancangan Aplikasi Prediksi Kelulusan Siswa Menggunakan Algoritma C4.5. *Progresif: Jurnal Ilmiah Komputer*, *16*(2), 121. https://doi.org/10.35889/progresif.v16i2.517

# LAMPIRAN

Adapun lampiran dari penelitian Penerapan Algoritma C4.5 Untuk Memprediksi Kelulusan Mahasiswa 2020 Fakultas Sains dan Teknologi (Studi Kasus: Teknologi Informasi, Teknik Lingkungan, Arsitektur, Kimia, Biologi) dapat diakses pada link berikut :

https://drive.google.com/drive/folders/19UmW8toYkZDHZNYVFMkYHB3Hosyw1PaE?usp=sharing