

**PREDIKSI KELULUSAN TEPAT WAKTU
PROGRAM PENDIDIKAN PROFESI DOKTER GIGI
MENGUNAKAN ALGORITMA *RANDOM FOREST***

LAPORAN TUGAS AKHIR

Laporan ini Disusun untuk Memenuhi Salah Satu Syarat Memperoleh Gelar Sarjana Strata 1 (S1) pada Program Studi Teknik Informatika Fakultas Teknologi Industri Universitas Islam Sultan Agung Semarang



DISUSUN OLEH :

DEDIE NUGROHO

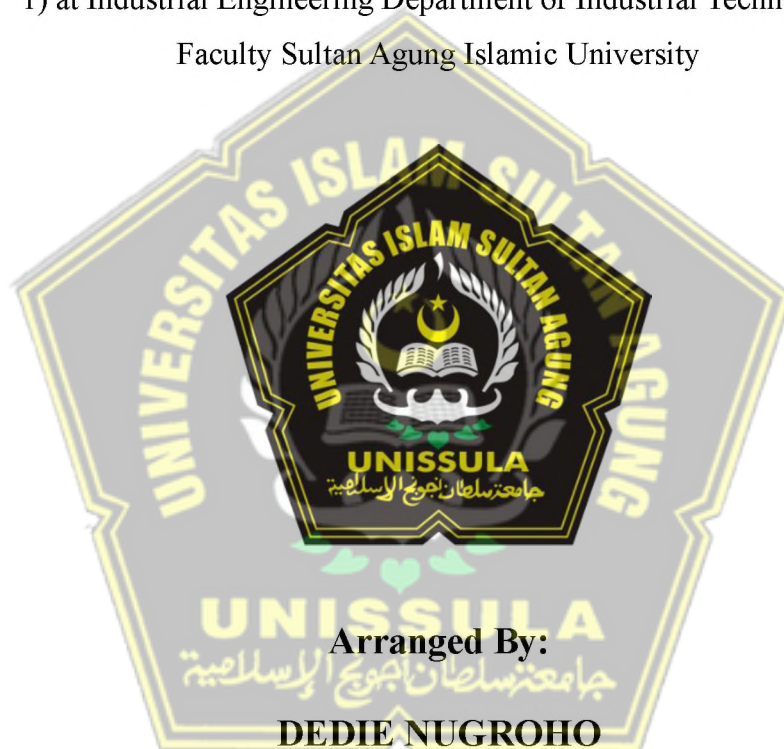
NIM 32602200151

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS ISLAM SULTAN AGUNG
SEMARANG
FEBRUARI 2026**

FINAL PROJECT

***PREDICTION OF ON-TIME GRADUATION IN
THE DENTIST PROFESSIONAL EDUCATION PROGRAM
USING THE RANDOM FOREST ALGORITHM***

Proposed to complete the requirement to obtain a bachelor 's degree (S-
1) at Industrial Engineering Department of Industrial Technology
Faculty Sultan Agung Islamic University



Arranged By:

DEDIE NUGROHO

NIM 32602200151

**MAJORING OF INFORMATICS ENGINEERING
INDUSTRIAL TECHNOLOGY FACULTY
SULTAN AGUNG ISLAMIC UNIVERSITY
SEMARANG
FEBRUARY 2026**

**LEMBAR PENGESAHAN
TUGAS AKHIR**

**PREDIKSI KELULUSAN TEPAT WAKTU PROGRAM
PENDIDIKAN PROFESI DOKTER GIGI MENGGUNAKAN
ALGORITMA *RANDOM FOREST***

**DEDIE NUGROHO
NIM 32602200151**

Telah dipertahankan di depan tim penguji ujian sarjana tugas akhir
Program Studi Teknik Informatika
Universitas Islam Sultan Agung
Pada tanggal : 12 Februari 2026

TIM PENGUJI UJIAN SARJANA :

Imam Much Ibnu Subroto, ST., M.Sc., Ph.D

NIK. 210600017

(Ketua Penguji)

20-02-2026

Sam Farisa Chaerul Haviana, ST., M.Kom

NIK. 210615046

(Anggota Penguji)

20-02-2026

Ghufron, ST, M.Kom

NIK. 210622056

(Pembimbing)

24-02-2026

Semarang, 12 Februari 2026

Mengetahui,

Kaprodi Teknik Informatika
Universitas Islam Sultan Agung

Moch Taufik, ST., MIT

MIDN. 210604034

SURAT PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertanda tangan dibawah ini :

Nama : Dedie Nugroho
NIM : 32602200151
Judul Tugas Akhir : PREDIKSI KELULUSAN TEPAT WAKTU
PROGRAM PENDIDIKAN PROFESI
DOKTER GIGI MENGGUNAKAN
ALGORITMA *RANDOM FOREST*

Dengan bahwa ini saya menyatakan bahwa judul dan isi Tugas Akhir yang saya buat dalam rangka menyelesaikan Pendidikan Strata Satu (S1) Teknik Informatika tersebut adalah asli dan belum pernah diangkat, ditulis ataupun dipublikasikan oleh siapapun baik keseluruhan maupun sebagian, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka, dan apabila di kemudian hari ternyata terbukti bahwa judul Tugas Akhir tersebut pernah diangkat, ditulis ataupun dipublikasikan, maka saya bersedia dikenakan sanksi akademis. Demikian surat pernyataan ini saya buat dengan sadar dan penuh tanggung jawab.

Semarang, 12 Februari 2026

Yang Menyatakan,



Dedie Nugroho

PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH

Saya yang bertanda tangan dibawah ini :

Nama : Dedie Nugroho
NIM : 32602200151
Program Studi : Teknik Informatika
Fakultas : Teknologi industri

Dengan ini menyatakan Karya Ilmiah berupa Tugas akhir dengan Judul : PREDIKSI KELULUSAN TEPAT WAKTU PROGRAM PENDIDIKAN PROFESI DOKTER GIGI MENGGUNAKAN ALGORITMA RANDOM FOREST. Menyetujui menjadi hak milik Universitas Islam Sultan Agung serta memberikan Hak bebas Royalti Non-Eksklusif untuk disimpan, dialihmediakan, dikelola dan pangkalan data dan dipublikasikan diinternet dan media lain untuk kepentingan akademis selama tetap menyantumkan nama penulis sebagai pemilik hak cipta. Pernyataan ini saya buat dengan sungguh-sungguh. Apabila dikemudian hari terbukti ada pelanggaran Hak Cipta/Plagiarisme dalam karya ilmiah ini, maka segala bentuk tuntutan hukum yang timbul akan saya tanggung secara pribadi tanpa melibatkan Universitas Islam Sultan agung.

Semarang, 12 Februari 2026

Yang menyatakan,



Dedie Nugroho

KATA PENGANTAR

Segala puji dan syukur penulis haturkan ke hadirat Allah SWT atas limpahan rahmat dan karunia-Nya, sehingga laporan tugas akhir dengan judul “Prediksi Kelulusan Tepat Waktu Program Pendidikan Profesi Dokter Gigi Menggunakan Algoritma *Random Forest*”. Keberhasilan penyusunan tugas akhir ini merupakan hasil dari dukungan banyak pihak. Penulis ingin menyampaikan rasa terima kasih yang mendalam kepada:

1. Ayah, Ibu, istriku dan anak-anakku yang telah memberikan semangat dan doa.
2. Bapak Prof. Dr. H. Gunarto, S.H., M.H selaku Rektor Universitas Islam Sultan Agung Semarang beserta seluruh jajaran.
3. Bapak Muhammad Qomaruddin, ST., M.Sc., Ph.D selaku Dekan Fakultas Teknologi Industri Universitas Islam Sultan Agung Semarang beserta seluruh jajaran.
4. Bapak Ghufron, ST, M.Kom yang senantiasa meluangkan waktu dan memberikan bimbingan dalam penulisan Tugas Akhir.
5. Seluruh Tenaga Pendidik dan Tenaga Kependidikan dilingkungan Program Studi Teknik Informatika Fakultas Teknologi Industri Universitas Islam Sultan Agung Semarang.
6. Dr. drg. Yayun Siti Rochmah, Sp.BM. dan seluruh civitas Fakultas Kedokteran Gigi Universitas Islam Sultan Agung Semarang.

Penulis menyadari sepenuhnya bahwa laporan tugas akhir ini masih jauh dari kata sempurna dan memiliki berbagai keterbatasan. Oleh karena itu, dengan segala kerendahan hati, penulis sangat terbuka terhadap saran dan kritik yang bersifat membangun guna penyempurnaan tulisan ini di masa mendatang.

Semarang, 12 Februari 2026

Penulis

Dedie Nugroho

DAFTAR ISI

LAPORAN TUGAS AKHIR	i
LEMBAR PENGESAHAN	iii
TUGAS AKHIR	iii
Sam Farisa Chaerul Haviana, ST., M.Kom	iii
Ghufron, ST, M.Kom	iii
Moch Taufik, ST., MIT	iii
SURAT PERNYATAAN KEASLIAN TUGAS AKHIR.....	iv
PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH.....	v
KATA PENGANTAR.....	vi
DAFTAR ISI	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN.....	xiii
ABSTRAK	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar belakang.....	1
1.2 Perumusan Masalah	3
1.3 Pembatasan Masalah	3
1.4 Tujuan.....	4
1.5 Manfaat	4
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA DAN DASAR TEORI	6

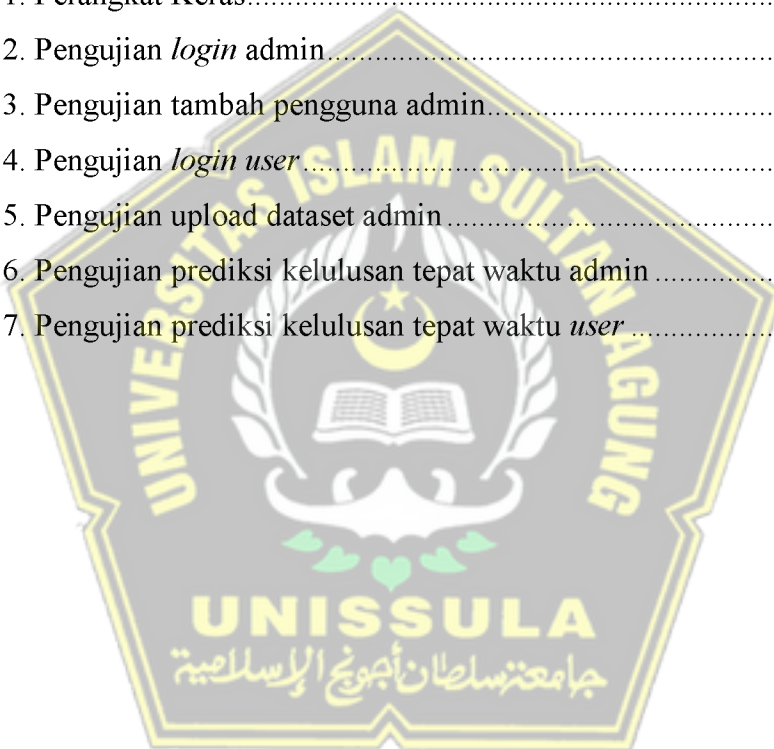
2.1. Tinjauan Pustaka	6
2.2. Dasar Teori	8
2.2.1. Prediksi.....	8
2.2.2. Data Mining	8
2.2.3. <i>Machine Learning</i>	9
2.2.4. Algoritma <i>Random Forest</i>	11
2.2.5. Evaluasi Model.....	12
BAB III METODE PENELITIAN.....	14
3.1. Studi Literatur	14
3.2. Pengumpulan Data	14
3.3. Perancangan Sistem Secara Umum.....	20
3.3.1. Proses Prediksi Kelulusan Tepat Waktu dengan <i>Random Forest</i> ..	20
3.3.1. Aktor	21
3.3.2. Alur kerja Sistem.....	21
3.4. Analisa Kebutuhan	23
3.4.1. <i>Use Case Diagram</i>	23
3.4.2. <i>Entity Relation Diagram (ERD)</i>	24
3.4.3. Perancangan Basis Data	25
3.4.4. <i>Data Flow Diagram (DFD)</i>	28
3.5. Perancangan Sistem	32
3.5.1. Perancangan Desain Sistem	32
BAB 4 HASIL DAN ANALISIS PENELITIAN.....	38
4.1. Implementasi perangkat lunak	38
4.2. Implementasi perangkat keras.....	38
4.3. Implementasi <i>User Interface</i>	39

4.4. Implementasi Training Model <i>Random Forest</i>	44
4.5. Implementasi Perhitungan Evaluasi model	47
4.6. Implementasi Random Forest Pada Prediksi Kelulusan Tepat Waktu ..	49
4.7. Pengujian Sistem	49
BAB 5 KESIMPULAN DAN SARAN	58
5.1. Kesimpulan	58
5.2. Saran	59
DAFTAR PUSTAKA	59



DAFTAR TABEL

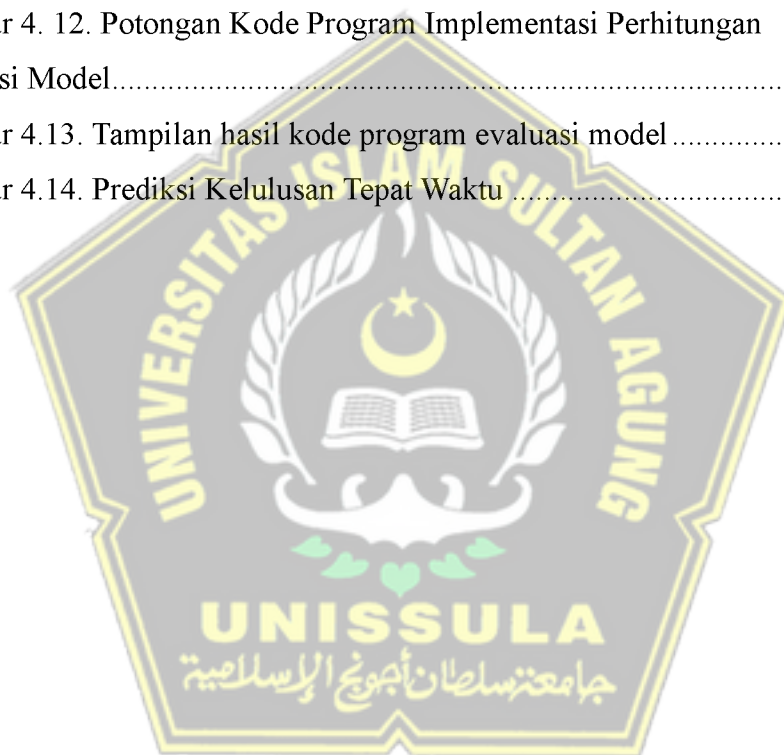
Tabel 3.1. pengguna	25
Tabel 3.2. dataset_mahasiswa	25
Tabel 3.3. data_preprocessing	26
Tabel 3.4. pelatihan_model	26
Tabel 3.5. simulasi_prediksi	27
Tabel 4.1. Perangkat Keras	38
Tabel 4.2. Pengujian <i>login</i> admin	50
Tabel 4.3. Pengujian tambah pengguna admin	50
Tabel 4.4. Pengujian <i>login user</i>	52
Tabel 4.5. Pengujian upload dataset admin	53
Tabel 4.6. Pengujian prediksi kelulusan tepat waktu admin	54
Tabel 4.7. Pengujian prediksi kelulusan tepat waktu <i>user</i>	56



DAFTAR GAMBAR

Gambar 3.1. Validasi tabel data_ukmppdg	16
Gambar 3.2. Validasi tabel_mahasiswa_lulus_profesi.....	16
Gambar 3.3. Validasi tabel mahasiswa_lulus_sarjana.....	17
Gambar 3.4. <i>Query</i> membuat dataset mahasiswa	18
Gambar 3. 5. Validasi hasil membuat dataset_mahasiswa.....	19
Gambar 3.6. Flowchart Alur Sistem.....	20
Gambar 3.7. Alur kerja sistem Admin.....	21
Gambar 3.8. Alur kerja Sistem <i>user</i>	22
Gambar 3.9. <i>Use Case Diagram</i>	23
Gambar 3.10. <i>Entity Relation Diagram</i> (ERD).....	24
Gambar 3.11. Diagram Konteks.....	28
Gambar 3.12. DFD Level 0.....	29
Gambar 3.13. DFD Level 1 Proses 1.....	30
Gambar 3.14. DFD Level 1 Proses 5.....	31
Gambar 3.15. DFD Level 1 Proses 7.....	31
Gambar 3.16. Halaman <i>login</i> aktor admin.....	32
Gambar 3.17. Halaman data pengguna.....	33
Gambar 3.18. <i>Form</i> input data pengguna.....	33
Gambar 3.19. Halaman dataset.....	34
Gambar 3.20. <i>Form Upload</i> dataset.....	34
Gambar 3.21. Halaman <i>Preprocessing</i>	35
Gambar 3.22. Halaman Training.....	35
Gambar 3.23. Halaman Simulasi Prediksi	36
Gambar 3.24. Halaman Simulasi Prediksi <i>User</i>	37
Gambar 4.1. <i>User Interface Login</i> admin dan user.....	39
Gambar 4.2. <i>User Interface</i> Data Pengguna	39
Gambar 4.3. <i>User Interface Form</i> Input Pengguna	40

Gambar 4.4. <i>User Interface</i> Dataset.....	41
Gambar 4. 5. <i>User Interface</i> Upload Dataset.....	41
Gambar 4. 6. <i>User Interface</i> Preprocessing.....	42
Gambar 4. 7. <i>User Interface</i> Training Model.....	42
Gambar 4.8. <i>User Interface</i> Simulasi Prediksi	43
Gambar 4.9. Potongan Kode Program Training Model	44
Gambar 4. 10. Potongan tampilan implementasi proses training model.....	45
Gambar 4. 11. Potongan Isi <i>File</i> Model Prediksi.....	46
Gambar 4. 12. Potongan Kode Program Implementasi Perhitungan Evaluasi Model.....	47
Gambar 4.13. Tampilan hasil kode program evaluasi model.....	48
Gambar 4.14. Prediksi Kelulusan Tepat Waktu	49



DAFTAR LAMPIRAN

Lampiran 1 Hasil Pelatihan Model dan Prediksi Kelulusan Tepat Waktu	63
Lampiran 2 Lembar Asistensi Pra Proposal Tugas Akhir	64
Lampiran 3 Lembar Revisi Proposal Tugas Akhir	65
Lampiran 4 Lembar Asistensi Pra Sidang Tugas Akhir	67



ABSTRAK

Jumlah Dokter Gigi per bulan Mei tahun 2025 di Indonesia masih kecil dibandingkan dengan jumlah penduduk. Pemerintah mengharapkan institusi Pendidikan tinggi sebagai pencetak lulusan Dokter Gigi diharapkan dapat meningkatkan jumlah lulusan Dokter Gigi. Hal itu tentu sangat penting bagi Program Pendidikan Profesi Dokter Gigi Fakultas Kedokteran Gigi Universitas Islam Sultan Agung bukan hanya untuk penilaian Akreditasi akan tetapi dapat berperan dalam meningkatkan jumlah dokter gigi Indonesia. Kegiatan bimbingan klinis dan evaluasi hasil pembelajaran klinis, Ujian Komprehensif telah dilakukan, namun dalam pelaksanaan ujian Uji Kompetensi Mahasiswa Program Pendidikan Profesi Dokter Gigi (UKMP2DG) yang telah dilakukan tidak semua berhasil lulus, maka penulis melakukan penelitian sebagai salah satu solusi inovatif adalah dengan memanfaatkan Machine Learning untuk melakukan prediksi kelulusan mahasiswa memanfaatkan data akademik serta algoritma Algoritma *Random Forest*. Data yang diambil memuat data peserta UKMP2DG Periode 1 Tahun 2020 sampai dengan periode 1 Tahun 2025. Hasil yang diperoleh dari penelitian yang telah dilaksanakan, penulis mendapatkan hasil evaluasi model dengan tingkat akurasi sebesar 83.19% , presisi 88.51% , recall 89.53% dan F1 Score 89.02% dengan fitur terbaik adalah IPK Sarjana dengan presentasi 41.57%.

Kata Kunci : Prediksi Kelulusan Mahasiswa, *Machine Learning*, *Random Forest*

ABSTRACT

As of May 2025, the number of dentists in Indonesia remains low relative to the population. In response, the government expects higher education institutions, as key producers of dental graduates, to help increase the supply of new dentists. This expectation is particularly significant for the Dental Professional Education Program at the Faculty of Dentistry, Sultan Agung Islamic University, which aims not only to meet accreditation standards but also to contribute meaningfully to addressing the national need for dentists. While clinical guidance, evaluation of clinical learning outcomes, and comprehensive examinations have been implemented, not all students have succeeded in passing the Dental Professional Education Program Student Competency Test (UKMP2DG). To address this challenge, Therefore, the author conducts research as one of the innovative solutions by utilizing Machine Learning to predict student graduation, leveraging academic data and the Random Forest algorithm. The dataset included UKMP2DG participants from Period 1 of 2020 through Period 1 of 2025. The results obtained from the conducted research show that the author achieved a model evaluation with an accuracy level of 83.19%, precision of 88.51%, recall of 89.53%, and F1 Score of 89.02%, with the best feature being the Bachelor's Degree GPA with a contribution of 41.57%.

Kata Kunci : *The Predict Student Graduation, Machine Learning, Random Forest*

BAB I

PENDAHULUAN

1.1 Latar belakang

Dokter gigi merupakan profesional medis dengan spesialisasi dalam bidang penyakit gigi dan mulut. Menurut data statistik Persatuan Dokter Gigi Indonesia (PDGI) per 15 Mei 2025 berjumlah 53.824 dokter gigi (PDGI, 2025). Jumlah itu terhitung sangat banyak, namun ternyata dibandingkan dengan jumlah populasi manusia di Indonesia sejumlah 284,44 juta jiwa (BPS Indonesia, 2025). Berdasarkan jumlah tersebut rasio antara dokter gigi dan jumlah penduduk hanya 0,189 dokter gigi per 1.000 penduduk, sedangkan rasio yang diharapkan adalah 1:3.000 padahal dari total 32 fakultas kedokteran gigi yang ada di Indonesia, hanya berhasil melahirkan 2.500 dokter gigi. Oleh sebab itu pemerintah melalui Menteri Kesehatan, mengharapkan bertambahnya lulusan dokter gigi di Indonesia untuk mencapai target tersebut dalam waktu yang tidak lama (Nadia Tarmizi, 2022).

Dalam mencapai jumlah dokter gigi tersebut tentunya institusi Pendidikan tinggi sebagai pencetak lulusan dokter gigi diharapkan dapat meningkatkan jumlah lulusan dan ini selaras dengan Peraturan Badan Akreditasi Nasional Perguruan Tinggi Nomor 27 Tahun 2024 tentang Instrumen Akreditasi Ulang Perguruan Tinggi, bahwa kelulusan tepat waktu menjadi salah satu elemen penilaian Akreditasi (Akademik, 2024). Dimana Mahasiswa bisa dinyatakan lulus jenjang profesi adalah mahasiswa yang telah menempuh 36 satuan kredit semester (SKS) dengan durasi kurikulum selama 2 semester, sebagaimana diatur dalam Peraturan Menteri Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia Nomor 53 Tahun 2023 mengenai Penjaminan Mutu Pendidikan Tinggi (Kemendikbudristek RI, 2023).

Hal itu tentu sangat penting bagi Fakultas Kedokteran Gigi (FKG) Program Pendidikan Profesi Dokter Gigi (PPDG) Universitas Islam Sultan Agung (UNISSULA) untuk penilaian Akreditasi, disamping itu juga dapat berperan dalam meningkatkan jumlah dokter gigi Indonesia. Oleh karena itu menjadi tugas

penting dalam mempersiapkan lulusan dokter gigi, bukan hanya dalam hal jumlah, namun juga kompeten dalam bidang kedokteran gigi sesuai ranah kedokteran gigi gigi yang selaras dengan Standar Kompetensi Dokter Gigi Indonesia (SKDGI). Untuk mencapai hal tersebut PPDG dalam proses kegiatan pembelajaran mahasiswa harus melaksanakan kegiatan bimbingan klinis dan evaluasi hasil pembelajaran klinis yang selanjutnya setelah menyelesaikannya mahasiswa dapat melakukan Ujian Komprehensif yang menilai kemampuan *knowledge*, *skill* dan *attitude* dengan metode ujian *Computer-Based Test* (CBT) dan *Objective Structured Clinical Examination* (OSCE). Setelah menyelesaikannya mahasiswa dapat mengikuti Uji Kompetensi Mahasiswa Program Pendidikan Profesi Dokter Gigi (UKMP2DG) dengan dua metode ujian yaitu Ujian Teori dan Ujian Praktek. Ujian ini sebagai evaluasi apakah mahasiswa dapat kompeten atau tidak kompeten sebagai dokter gigi dengan syarat harus lulus UKMP2DG (Program Profesi Dokter Gigi, 2024).

Namun dalam pelaksanaan ujian UKMP2DG yang telah dilakukan mahasiswa PPDG FKG UNISSULA tidak semua berhasil lulus. Tidak sedikit mahasiswa yang mendapatkan hasil tidak lulus menurut data tahun 2023, 20% Tidak lulus dalam UKMP2DG. Tentunya ini menjadi tantangan bagi PPDG melihat tingginya angka kegagalan mahasiswa dalam ujian tersebut karena akan menambah masa studi mahasiswa yang berakibat pada capaian kelulusan tepat waktu. Fenomena ini memerlukan pendekatan preventif untuk memitigasi risiko ketidاكلulusan sejak dini. Salah satu solusi inovatif adalah dengan memanfaatkan *Machine Learning* untuk meramalkan lulus atau tidaknya mahasiswa melalui rekam jejak akademik masa lalu dengan Algoritma *Random Forest* karena kemampuannya mengolah data akademik seperti ipk dengan Tingkat akurasi 87,5% (Andrianof dkk., 2025).

Penulis menunjukkan minat kuat untuk melaksanakan penelitian berdasarkan penjelasan latar belakang yang telah dipaparkan dengan pokok utama Prediksi Kelulusan Mahasiswa dengan menggunakan algoritma *Random Forest*. Hasil prediksi kelulusan diharapkan menjadi gambaran awal bagi civitas akademik PPDG

FKG UNISSULA dalam mengoptimalkan strategi pembelajaran, memberikan bimbingan khusus kepada mahasiswa berisiko gagal, dan akhirnya meningkatkan persentase kelulusan UKMP2DG kemudian pada akhirnya berpengaruh pada hasil kelulusan tepat waktu mahasiswa PPDG.

1.2 Perumusan Masalah

Dari penjelasan latar belakang masalah tersebut, dirumuskan pertanyaan penelitian sebagai berikut:

1. Bagaimana cara menggunakan data akademik untuk memprediksi mahasiswa PPDG FKG UNISSULA yang berisiko tidak lulus UKMP2DG sejak dini?
2. Bagaimana membangun model prediksi risiko ketidakkelulusan UKMP2DG menggunakan algoritma *Random Forest* yang mampu mengidentifikasi mahasiswa berisiko tidak lulus sejak dini ?

1.3 Pembatasan Masalah

Dari latar belakang dan rumusan masalah yang telah diuraikan, pembatasan ruang lingkup penelitian ini mencakup hal-hal berikut :

1. Penelitian dilakukan di Program Pendidikan Profesi Dokter Gigi Fakultas Kedokteran Gigi Universitas Islam Sultan Agung Semarang.
2. Data yang diambil memuat data peserta UKMP2DG Periode 1 Tahun 2020 sampai dengan periode 1 Tahun 2025.

1.4 Tujuan

Tujuan dari penelitian ini adalah

1. Mengeksplorasi secara kritis dan mendeteksi secara spesifik faktor-faktor data akademik mahasiswa PPDG FKG UNISSULA yang paling berpengaruh terhadap kelulusan UKMP2DG.
2. Mengembangkan model prediksi dengan algoritma *Random Forest* berbasis data akademik yang mampu mengidentifikasi mahasiswa berisiko tidak lulus sejak dini serta mengevaluasi akurasi model prediksi.
3. Merancang sistem prediksi berbasis website yang mengintegrasikan model *Random Forest* untuk mengetahui mahasiswa yang berisiko tidak lulus UKMP2DG sebagai dasar pemberian *treatment* khusus oleh pihak PPDG FKG UNISSULA.

1.5 Manfaat

Keuntungan pokok yang diperoleh penulis melalui pelaksanaan penelitian ini meliputi pemahaman mendalam terhadap tahapan-tahapan yang dilakukan dalam prediksi kelulusan tepat waktu mahasiswa, mengetahui algoritma yang dipelajari, persentase tingkat akurasi dalam prediksi serta tingkat kesulitan dalam membangun aplikasi.

1.6 Sistematika Penulisan

Struktur penulisan yang diterapkan penulis dalam merangkai laporan tugas akhir ini disusun sebagai berikut:

BAB 1: PENDAHULUAN

Pada bagian ini, penulis memaparkan alasan pemilihan topik, pertanyaan penelitian, pembatasan ruang lingkup,

sasaran studi, pendekatan metodologi, serta kerangka penyusunan tulisan.

BAB 2: TINJAUAN PUSTAKA DAN DASAR TEORI

Bab ini memuat sejumlah landasan teori yang menjadi rujukan utama dalam tahap perancangan dan pengembangan sistem, tinjauan studi terdahulu, kerangka pemikiran, penyelesaian model matematis, pengembangan sistem, serta desain pengujian.

BAB 3: METODE PENELITIAN

Bab ini menguraikan strategi penelitian yang terdiri dari tahapan sistematis untuk menangani isu utama, penggalian data, dan proses pembentukan model.

BAB 4: HASIL PENELITIAN

Bab ini mengemukakan output dari tahap perancangan dan pembangunan sistem. Di samping itu, bab ini turut memuat hasil uji coba sistem melalui metode black box testing yang diterapkan pada penelitian ini.

BAB 5: KESIMPULAN DAN SARAN

Bab ini menguraikan kesimpulan dari analisis hasil penelitian beserta rekomendasi yang diajukan.

BAB II

TINJAUAN PUSTAKA DAN DASAR TEORI

2.1. Tinjauan Pustaka

Penelitian mengenai prediksi kelulusan mahasiswa berbasis ML khususnya menggunakan Algoritma *Random Forest* sudah dilaksanakan oleh peneliti sebelumnya. Dalam studi yang telah kerjakan oleh (Andrianof *dkk.*, 2025) mendapatkan hasil penelitian yang menunjukkan bahwa algoritma *Random Forest* memiliki akurasi sebesar 87.5 % dengan variabel IPK dan kehadiran sebagai faktor paling penting dalam menentukan pada waktunya. Sementara studi lainnya melakukan ramalan kelulusan mahasiswa melalui algoritma *Random Forest* dan K-NN berbasis PSO dan hasil yang diperoleh adalah *Random Forest* memiliki akurasi yang lebih akurat sebesar 95,79% dibandingkan algoritma K-NN sebesar 93,49% dengan atribut yang digunakan Nomor Pokok Mahasiswa (NPM), identitas lengkap, program studi, fakultas, tahun masuk perkuliahan, serta tahun lulus (Indra *dkk.*, 2023).

Pada konferensi internasional juga telah menunjukkan temuan bahwa memanfaatkan algoritma *Random Forest* menghasilkan nilai akurasi yang baik 96.88% (Ghosh dan Janan, 2021). Begitu juga yang dilakukan oleh Balabied dan Eid pada tahun 2023, menunjukkan bahwa algoritma *Random Forest* mendapatkan hasil akurasi mencapai 90% dalam mendeteksi siswa yang beresiko ketertinggalan akademik pada lingkungan pembelajaran terbuka (Balabied dan Eid, 2023).

Peneliti lain menggabungkan data akademik dan non akademik seperti Indeks Prestasi Sementara (IPS) dari semester 1 sampai 4, kunjungan, pinjam buku, dan status. Penelitian ini juga membandingkan dua algoritma dengan hasil algoritma *Random Forest* mencapai tingkat akurasi yang lebih unggul dibandingkan C.45, yakni sebesar 95,13% (Ubaidilah, 2023).

Penelitian dalam pengimplementasian algoritma *Random Forest* oleh (Nurfadilla dan Faisal, 2022) juga mendapatkan hasil nilai akurasi yang besar yaitu sebesar 90.74% dengan variabel yang digunakan adalah IPS. Bahkan penelitian yang telah dilakukan mendapatkan hasil akurasi 100% dalam prediksi mahasiswa lulus tepat waktu dengan variabel gender, usia, status perkawinan, dan pekerjaan (Riskiyono dan Mahdiana, 2024).

(Putra dan Harahap, 2024) dalam penelitian yang dilakukan menghasilkan Tingkat prediksi kelulusan mahasiswa sebesar 88% menggunakan variabel IP semester ke-5, IP semester ke-4, IP semester ke-3, IP semester ke-2, IP semester ke-1, jenis kelamin, dan status pernikahan. Kemudian (Diantika *dkk.*, 2024) mendapatkan hasil yang cukup signifikan dengan akurasi 90,04% menggunakan data variabel Indeks Prestasi Kumulatif (IPK) mahasiswa. Penelitian terbaru tentang prediksi kelulusan mahasiswa yang telah dilakukan oleh (Ermatita, 2025) Penelitian lain menunjukkan akurasi sebesar 98,41% dengan variabel yang digunakan adalah IPK, SKS.

Berdasarkan penelitian terdahulu, algoritma *Random Forest* terbukti menghasilkan akurasi tinggi dalam meramalkan keberhasilan kelulusan mahasiswa berdasarkan catatan akademik seperti IPK, IPS, SKS, kehadiran maupun data non akademik seperti status pekerjaan, status pernikahan dan kegiatan perpustakaan. Namun demikian, masih terdapat kesempatan untuk menciptakan model prediksi yang lebih menyeluruh dengan memperhatikan berbagai jenis variabel serta konteks lembaga pendidikan yang bervariasi. Karenanya, penelitian ini akan menerapkan algoritma *Random Forest* untuk memperkirakan kelulusan mahasiswa PPDG FKG UNISSULA dalam jangka waktu yang ditentukan dengan kombinasi variabel-variabel penting. Penelitian ini diharapkan mampu menciptakan model prediksi yang tepat, bermanfaat, dan berfungsi sebagai alat bantu untuk keputusan akademis PPDG FKG UNISSULA.

2.2. Dasar Teori

2.2.1. Prediksi

Berdasarkan definisi dalam Kamus Besar Bahasa Indonesia (KBBI), prediksi dapat dipahami sebagai sebuah kesimpulan atau gambaran mengenai hal-hal yang akan datang, yang diperoleh melalui proses meramal atau memperkirakan dengan berlandaskan pada data dan pola dari masa lampau. Pada dasarnya prediksi seperti cara sistematis kita menebak apa yang akan terjadi nanti, dengan mengandalkan informasi dan pola dari pengalaman sebelumnya hingga sekarang yang bertujuan memperkecil kesalahan antara hasil yang diperkirakan dengan kejadian sebenarnya (Diantika *dkk.*, 2024).

Prediksi memainkan peran krusial dalam ranah akademik yang akan yang akan memberikan dampak yang signifikan, dengan Model prediktif ini memungkinkan lembaga pendidikan untuk menyusun intervensi yang presisi, mengoptimalkan alokasi sumber daya pendukung, mempertinggi angka kelulusan on-time, serta memberikan bantuan yang lebih tepat guna bagi siswa berpotensi berisiko. (Andrianof *dkk.*, 2025). Dari pengertian tersebut maka prediksi kelulusan mahasiswa ini bisa menjadi bahan evaluasi dari data lulusan sebelumnya yang hasilnya akan berguna dalam pengambilan keputusan pada PPDG FKG UNISSULA

2.2.2. Data Mining

Data mining dapat dipahami sebagai sebuah sintesis keilmuan, yakni titik temu antara berbagai disiplin seperti pembelajaran mesin (*machine learning*), penggalian pola, ilmu statistika, pengelolaan basis data, dan metode visualisasi yang dapat menyelesaikan tantangan dalam mengekstraksi informasi berharga dari himpunan banyak data yang terhimpun rapi dalam sebuah sistem basis data (Indra *dkk.*, 2023). data mining dapat diibaratkan sebagai aktivitas menambang dalam dunia data. Ia berupaya menemukan butiran-butiran informasi dan pola tersembunyi dari timbunan data yang sangat besar, melalui proses pengumpulan,

ekstraksi, analisis, dan perhitungan statistika. Dalam percakapan sehari-hari di dunia teknologi, proses ini sering disebut sebagai *knowledge discovery*, *knowledge extraction*, *data/pattern analysis*, atau *information harvesting*. Dari proses tersebut maka akan menghasilkan pengetahuan yang sangat bermanfaat. Data mining dibagi menjadi beberapa kategori berdasarkan jenis tugasnya, yaitu deskripsi, estimasi, prediksi, klasifikasi, pengelompokan, serta asosiasi. Khususnya kelompok prediksi yang memiliki keterkaitan erat dengan dua saudara dekatnya, yaitu klasifikasi serta estimasi. Ketiganya sama-sama bergerak di ranah peramalan, di mana prediksi secara khusus bertugas untuk meramalkan angka masa depan melalui berbagai metode serta teknik yang mirip dengan keduanya. (Amna dkk., 2023).

Data mining dapat meringankan beban kerja, serta efektif digunakan dalam berbagai bidang. Sebagai contoh dalam bidang pemasaran, pengelola bisnis dapat mengetahui pola tren tersembunyi pada transaksi yang dilakukan pelanggan. Pada bidang perbankan, teknik ini dapat membantu penipuan kartu kredit dan mengetahui loyalitas dari nasabah. Dalam bidang kedokteran dapat membantu dokter dalam mengidentifikasi pola dari terapi medis yang sukses untuk penyakit. Bidang manufaktur juga dapat memanfaatkan data mining untuk mendeteksi kesalahan dan menentukan parameter kontrol yang ideal. Dalam bidang pendidikan efektif dimanfaatkan untuk meramalkan probabilitas kelulusan mahasiswa sesuai dengan target waktu yang ditentukan yang memudahkan pengklasifikasian antara yang lulus on-time atau tidak (Nurfadilla dan Faisal, 2022). Kemahiran ini menjadi fondasi utama data mining dalam penelitian ini, guna mengungkap pola-pola tersembunyi yang mampu meramalkan apakah seorang mahasiswa berpotensi menyelesaikan studi tepat waktu atau tidak pada PPDG FKG UNISSULA. Hasil identifikasi tersebut dapat menjadi acuan untuk menentukan siapa saja mahasiswa yang harus mendapatkan perlakuan khusus.

2.2.3. Machine Learning

Secara konseptual, *Machine Learning (ML)* merupakan cabang *Artificial Intelligence (AI)* yang memungkinkan sistem komputer untuk melakukan proses

belajar mandiri berdasarkan data, sehingga ia dapat menemukan pola dan pengetahuan tanpa harus diprogram secara eksplisit. Dengan menggunakan algoritma yang telah melalui proses pelatihan data, model ML mampu melakukan prediksi dan pengelompokan terhadap data-data baru. ML mampu memproses dan menganalisis data dalam skala besar serta kompleks dengan efektif, dengan demikian mendukung pengambilan keputusan serta prediksi yang lebih tepat dan akurat di berbagai sektor (Habibie Sukarna *dkk.*, 2024). ML terbagi menjadi 3 kelompok *Supervised Learning*, *Unsupervised Learning*, *Reinforcement Learning* (Nurhalizah *dkk.*, 2024).

Supervised Learning Merupakan model ML dalam Proses pembelajaran berbasis data berlabel ini memungkinkan model untuk meramalkan dan mengklasifikasikan data baru, dengan mengandalkan pola-pola yang telah dipelajari sebelumnya. *Supervised Learning* pada dasarnya terdiri dari dua elemen kunci: input berupa fitur-fitur, serta output berupa label kelas. Model akan menganalisis input dan output selama proses pelatihan, kemudian setelah model terbentuk dilakukan proses pengujian untuk mengukur performa model dalam meramalkan data yang belum pernah ditemui sebelumnya. *Supervised Learning* pada umumnya digunakan pada untuk menyelesaikan masalah klasifikasi dan regresi. Dalam implementasinya, algoritma *Supervised Learning* seperti algoritma tetangga terdekat K (KNN), Bayes Sederhana, Pohon Keputusan, serta Mesin Vektor Penyangga (SVM) telah banyak diterapkan pada penelitian karena kemampuannya dalam menemukan pola berlabel serta menyajikan hasil akurasi yang baik. Meskipun memiliki akurasi yang baik namun model sangat dipengaruhi oleh kualitas data. Selain itu, model sangat mudah kehilangan kemampuan generalisasi karena terlalu fokus pada data latih bila rancang bangun model menjadi berlebihan rumit atau kumpulan data latihannya tidak memadai (Nurhalizah *dkk.*, 2024).

Penelitian ini secara spesifik menerapkan model *Supervised Learning* untuk klasifikasi dengan algoritma Random Forest yang menggabungkan model Decision Tree untuk mendapatkan hasil prediksi yang akurat. Seperti penelitian yang telah

dilakukan dimana Model mampu Membaca pola kelulusan mahasiswa melalui pendekatan prediktif dengan tingkat akurasi optimal (Andrianof *dkk.*, 2025).

2.2.4. Algoritma *Random Forest*

Algoritma *Random Forest* adalah salah algoritma ML terdiri dari kombinasi pohon klasifikasi yang bekerja secara mandiri. Hasil prediksi diperoleh dari voting mayoritas. Data yang akan dipakai kemudian diklasifikasikan. Pada penelitian ini data terbagi menjadi dua kategori lulus sesuai jadwal dan tidak lulus sesuai jadwal (Nurfadilla dan Faisal, 2022). Berikut cara kerja dari algoritma *Random Forest* :

1. Himpunan data D yang mencakup n sampel dipartisi secara tidak berurutan menjadi m bagian, membentuk D_i , di mana setiap bagian D_i dimanfaatkan untuk membangun satu pohon keputusan sesuai dengan rumus berikut:

$$D = \{X_1, Y_1\}, X_1 \sim D \quad (1)$$

2. Selanjutnya *Decision Tree* dibangun pada setiap cabang simpul D_i , melalui seleksi ciri secara arbitrary dari seluruh atribut yang ada. Jika terdapat k atribut, maka hanya atribut paling optimal yang dipilih berdasarkan metrik *Gini Index*

$$G = \sum_{i=1}^c p_i^2 \quad (2)$$

3. Setelah *Decision tree* dibuat kemudian menggabungkan hasil dari *Decision tree* dengan menggunakan voting mayoritas untuk menentukan kelas final dengan perhitungan sebagai berikut :

$$\hat{Y} = \text{mode}(\{T_1(x), T_2(x), T_m(x)\}) \quad (3)$$

T_i adalah prediksi kelas dari pohon ke-1 untuk input nilai x .

4. Selanjutnya adalah model dievaluasi melalui indikator akurasi, tingkat presisi, *recall*, serta *F1-score* (Marwa Sulehu dan First Wanita, 2025).

2.2.5. Evaluasi Model

Dengan *confusion matrix* kita akan mendapatkan matrik evaluasi akurasi, presisi, *recall*, dan *F1-score* (Diantika *dkk.*, 2024).

1. Akurasi

Dalam model prediksi kelulusan, akurasi digunakan oleh sebagai indikator evaluasi algoritma *Random Forest* dalam prediksi kelulusan mahasiswa tepat waktu (Diantika *dkk.*, 2024). Berikut adalah rumus perhitungan akurasi (Dutt, 2018).

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

Keterangan :

TP = *True Positive*

TN = *True Negative*

FP = *False Positive*

FN = *False Negative*

Dalam model prediksi kelulusan, akurasi digunakan oleh sebagai indikator evaluasi algoritma *Random Forest* dalam prediksi kelulusan mahasiswa tepat waktu (Diantika *dkk.*, 2024).

2. Presisi

Presisi menilai ketepatan model dalam mendeteksi kasus positif secara akurat, yakni melalui perbandingan antara prediksi positif yang benar dengan total prediksi positif yang dihasilkan. Presisi sangat penting dilakukan untuk mengetahui seberapa presisi prediksi positif benar-benar relevan (Dutt, 2018).

$$Presisi = \frac{TP}{TP + FP} \quad (5)$$

Presisi dalam prediksi kelulusan digunakan mengukur ketepatan prediksi positif untuk kelulusan tepat waktu yang dihasilkan model (Andrianof *dkk.*, 2025).

3. *Recall*

Recall dimanfaatkan untuk menentukan seberapa baik mendapatkan seluruh prediksi positif yang sebenarnya dari semua prediksi.

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

Dalam prediksi kelulusan digunakan untuk membandingkan ketepatan lulus mahasiswa berdasarkan kelas tersebut dengan kelas ketepatan waktu lulus (Diantika *dkk.*, 2024)

4. *F1-Score*

Mengintegrasikan presisi dan *recall* dengan mengambil mean harmonik dari presisi dan recall untuk menilai kinerja model dengan rumus persamaan sebagai berikut:

$$F1 - Score = \frac{2 \times Presisi \times Recall}{Presisi + Recall} \quad (7)$$

Dalam konteks prediksi kelulusan mahasiswa, F1-score berfungsi untuk mengukur keseimbangan performa model antar dua kelas, yaitu peserta didik menyelesaikan studi sesuai jadwal dan yang mengalami keterlambatan kelulusan (Nurfadilla dan Faisal, 2022).

BAB III

METODE PENELITIAN

3.1. Studi Literatur

Dalam tahap ini penulis melakukan studi literatur sebagai rujukan seperti jurnal ilmiah, skripsi, situs, serta dokumentasi lainnya yang berhubungan dengan penelitian ini. Studi ini bertujuan untuk memahami algoritma *Random Forest*, evaluasi model serta teknologi pengembangan sistem prediksi. Studi literatur yang dipelajari meliputi prediksi, *Data Mining*, *Machine Learning*, algoritma *Random Forest*, serta penerapannya dalam sistem berbasis web.

3.2. Pengumpulan Data

Taha ini secara bertahap menyusun kumpulan data esensial untuk penelitian dengan mengakses sumber dari divisi akademik, memastikan kesesuaian jenis data yang dibutuhkan PPDG dan Program Studi Sarjana Kedokteran Gigi (PSKG) di FKG UNISSULA yang meliputi tiga jenis data yaitu

1. Data UKMP2DG

Data ini berisi informasi data hasil UKMP2DG periode 1 tahun 2020 sampai dengan periode 1 tahun 2025 yang berjumlah 564 baris data. Data mencakup nomor peserta ujian, nama mahasiswa PPDG, periode ujian UKMP2DG, tanggal ujian UKMP2DG dan kompetensi yang berisi keterangan hasil ujian UKMP2DG.

2. Data kelulusan mahasiswa PPDG

Data ini berisi informasi data akademik kelulusan mahasiswa PPDG yang telah menempuh pendidikan PPDD dari kelulusan Periode 1 Tahun 2014 sampai dengan periode 43 tahun 2025 yang berjumlah 616 baris data yang terdiri dari nomor induk mahasiswa PPDG, nama mahasiswa PPDG, dan IPK PPDG.

3. Data kelulusan mahasiswa PSKG

Data ini berisi informasi data akademik kelulusan mahasiswa PPDG yang telah menempuh pendidikan PSKG dari kelulusan periode 1 tahun 2013 sampai sampai dengan periode 27 tahun 2025 yang berjumlah 981 baris data. Data tersebut terdiri dari nomor induk mahasiswa PSKG , nama mahasiswa PSKG, jurusan, tahun masuk, IPK mahasiswa PSKG, jenis keluar, dan status kelulusan mahasiswa PSKG

Seluruh data yang diberikan berformat *Microsoft Excel* (.xlsx). Data yang telah dikumpulkan disimpan dalam folder untuk dilakukan verifikasi kelengkapan data. Proses selanjutnya data di konversi dari format .xlsx ke dalam format .csv dengan cara melakukan *save as* dan memilih CVS (*Comma delimited*) (*.csv). Setelah seluruh file dikonversi selanjutnya dilakukan migrasi kedalam *Database Management System* (DBMS) Mysql melalui antarmuka phpMyAdmin untuk memudahkan dalam proses pengolahan data dengan urutan tahapan berikut::

1. Membentuk basis data baru bernama koleksi_data
2. Membuat tabel tiga tabel yang terdiri dari data_ukmppdg untuk menyimpan data peserta UKMP2DG, tabel mahasiswa_profesi untuk menyimpan data kelulusan PPDG dan tabel mahasiswa_sarjana untuk menyimpan hasil data kelulusan SKG.
3. Memasukkan data ke masing-masing tabel yang telah dibuat dengan menggunakan fitur *import wizard* pada phpMyAdmin.
4. Melaksanakan verifikasi data guna memastikan semua data yang telah diinput telah sesuai dengan data sebenarnya.

Your SQL query has been executed successfully.

```
SELECT 'data_ukmppdg' AS nama_tabel, (SELECT GROUP_CONCAT(column_name ORDER BY
ordinal_position SEPARATOR ', ') FROM information_schema.columns WHERE table_schema =
'koleksi_data' AND table_name = 'data_ukmppdg') AS nama_kolom, (SELECT COUNT(*) FROM
information_schema.columns WHERE table_schema = 'koleksi_data' AND table_name =
'data_ukmppdg') AS jumlah_kolom, (SELECT COUNT(*) FROM data_ukmppdg) AS jumlah_baris;
```

Profiling [[Edit inline](#)] [[Edit](#)] [[Explain SQL](#)] [[Create PHP code](#)] [[Refresh](#)]

Extra options

nama_tabel	nama_kolom	jumlah_kolom	jumlah_baris
data_ukmppdg	no_pst, nama_mhs, periode, tanggal, kompetensi	5	564

Gambar 3.1. Validasi tabel data_ukmppdg

Pada gambar 3.1. Validasi tabel data_ukmppdg adalah proses melakukan validasi data dari hasil masukkan data UKMP2DG dengan hasil bahwa data yang dimasukkan sudah sesuai dengan nama tabel data_ukmppdg dengan jumlah kolom 5 yang terdiri dari no_pst (nomor peserta ujian UKMP2DG), nama_mhs (nama mahasiswa PPDG), periode (periode ujian UKMP2DG), tanggal (tanggal ujian UKMP2DG), kompetensi. Jumlah data pada tabel berjumlah 564 baris, artinya data hasil import dama dengan data UKMP2DG.

Your SQL query has been executed successfully.

```
SELECT 'mahasiswa_lulus_profesi' AS nama_tabel, (SELECT GROUP_CONCAT(column_name ORDER
BY ordinal_position SEPARATOR ', ') FROM information_schema.columns WHERE table_schema
= 'koleksi_data' AND table_name = 'mahasiswa_lulus_profesi') AS nama_kolom, (SELECT
COUNT(*) FROM information_schema.columns WHERE table_schema = 'koleksi_data' AND
table_name = 'mahasiswa_lulus_profesi') AS jumlah_kolom, (SELECT COUNT(*) FROM
mahasiswa_lulus_profesi) AS jumlah_baris;
```


Profiling [[Edit inline](#)] [[Edit](#)] [[Explain SQL](#)] [[Create PHP code](#)] [[Refresh](#)]

Extra options

nama_tabel	nama_kolom	jumlah_kolom	jumlah_baris
mahasiswa_lulus_profesi	nim_pro, nama_pro, ipk_pro	3	616

Gambar 3.2. Validasi tabel_mahasiswa_lulus_profesi

Pada gambar 3.2. Validasi tabel `mahasiswa_lulus_profesi` merupakan validasi data dari hasil import data kelulusan mahasiswa PPDG yang dimasukkan kedalam tabel `mahasiswa_lulus_profesi` dengan jumlah kolom 3 yang terdiri dari `nim_pro` (nomor induk mahasiswa PPDG), `nama_pro` (nama mahasiswa PPDG) dan `ipk_pro` (IPK mahasiswa PPDG). Jumlah baris pada tabel tersebut berjumlah 616 baris, artinya bahwa hasilnya sesuai dengan data kelulusan mahasiswa PPDG.



Your SQL query has been executed successfully.

```

SELECT 'mahasiswa_lulus_sarjana' AS nama_tabel, (SELECT GROUP_CONCAT(column_name
ORDER BY ordinal_position SEPARATOR ', ') FROM information_schema.columns WHERE
table_schema = 'koleksi_data' AND table_name = 'mahasiswa_lulus_sarjana') AS
nama_kolom, (SELECT COUNT(*) FROM information_schema.columns WHERE table_schema =
'koleksi_data' AND table_name = 'mahasiswa_lulus_sarjana') AS jumlah_kolom, (SELECT
COUNT(*) FROM mahasiswa_lulus_sarjana) AS jumlah_baris;

```

Profiling [[Edit inline](#)] [[Edit](#)] [[Explain SQL](#)] [[Create PHP code](#)] [[Refresh](#)]

Extra options

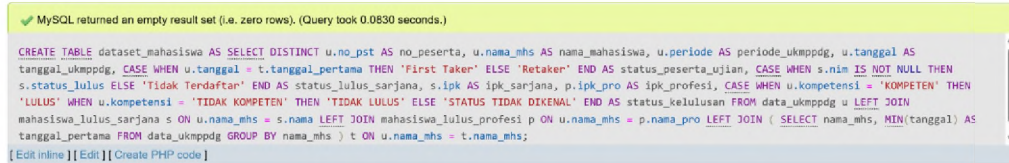
nama_tabel	nama_kolom	jumlah_kolom	jumlah_baris
mahasiswa_lulus_sarjana	nim, nama, jurusan, tahun_masuk, ipk, jenis_keluar, status_lulus	7	981

Gambar 3.3. Validasi tabel `mahasiswa_lulus_sarjana`

Pada gambar 3.3. Validasi tabel `mahasiswa_lulus_sarjana` merupakan hasil dari import data kelulusan mahasiswa PSKG ke dalam tabel `mahasiswa_lulus_sarjana` dengan jumlah kolom 7 yang mencakup `nim` (nomor induk mahasiswa PSKG), `nama` (nama mahasiswa PSKG), `jurusan`, `tahun_masuk`, `ipk` (IPK mahasiswa PSKG), `jenis_keluar`, `status_lulus` (status kelulusan mahasiswa PSKG). Jumlah baris hasil validasi berjumlah 981, maka berdasarkan hasil tersebut data hasil import yang masuk ke dalam tabel sesuai dengan data kelulusan PSKG.

5. Data yang sudah tersimpan dalam database dilakukan integrasi untuk membentuk tabel `dataset_mahasiswa` dengan melakukan *query* untuk menggabungkan ketiga tabel data dengan nama mahasiswa sebagai penghubung serta hanya mengambil data nama mahasiswa PPDG pada tabel `data_ukmppdg`. Selain

menggabungkan dalam tahap ini dilakukan proses transformasi label agar lebih mudah dipahami.



```

MySQL returned an empty result set (i.e. zero rows). (Query took 0.0830 seconds.)

CREATE TABLE dataset_mahasiswa AS SELECT DISTINCT u.no_pst AS no_peserta, u.nama_mhs AS nama_mahasiswa, u.periode AS periode_ukmppdg, u.tanggal AS tanggal_ukmppdg, CASE WHEN u.tanggal = t.tanggal_pertama THEN 'First Taker' ELSE 'Retaker' END AS status_peserta_ujian, CASE WHEN s.nim IS NOT NULL THEN s.status_lulus ELSE 'Tidak Teraftar' END AS status_lulus_sarjana, s.ipk AS ipk_sarjana, p.ipk_pro AS ipk_profesi, CASE WHEN u.kompetensi = 'KOMPETEN' THEN 'LULUS' WHEN u.kompetensi = 'TIDAK KOMPETEN' THEN 'TIDAK LULUS' ELSE 'STATUS TIDAK DIKENAL' END AS status_kelulusan FROM data_ukmppdg u LEFT JOIN mahasiswa_lulus_sarjana s ON u.nama_mhs = s.nama LEFT JOIN mahasiswa_lulus_profesi p ON u.nama_mhs = p.nama_pro LEFT JOIN ( SELECT nama_mhs, MIN(tanggal) AS tanggal_pertama FROM data_ukmppdg GROUP BY nama_mhs ) t ON u.nama_mhs = t.nama_mhs;
[Edit inline] [Edit] [Create PHP code]

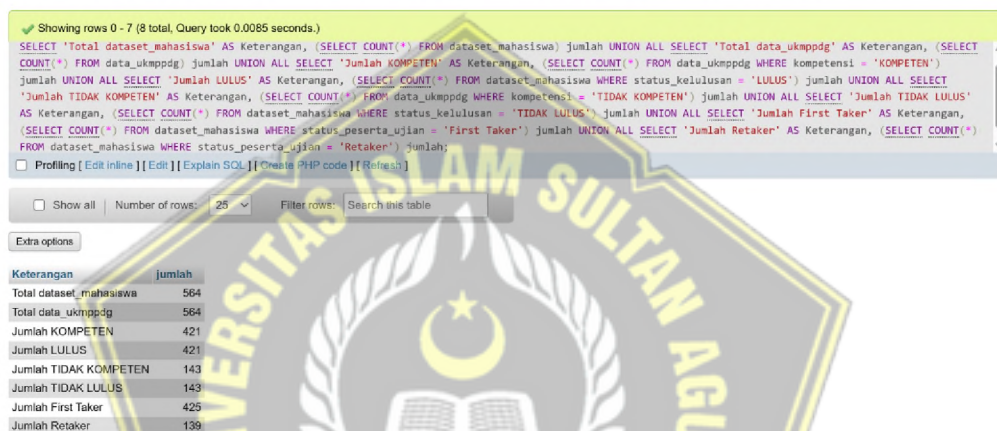
```

Gambar 3.4. Query membuat dataset mahasiswa

Pada gambar 3.4. Query membuat dataset mahasiswa merupakan proses pembuatan tabel dataset_mahasiswa dengan menggabungkan tiga tabel yang terdiri dari data_ukmppdg, mahasiswa_lulus_sarjana, dan mahasiswa_lulus_profesi. Tabel utama yang digunakan adalah data_ukmppdg. Query tersebut melakukan pembuatan tabel dataset_mahasiswa untuk menampung data dari query penggabungan tiga tabel tersebut dengan nama mahasiswa PPDG sebagai kunci penghubung dengan hasil sebagai berikut:

- a. no_peserta berisi nomor peserta ujian UKMP2DG dari tabel data_ukmppdg kolom no_pst.
- b. nama_mahasiswa merupakan data nama mahasiswa PPDG dari tabel data_ukmppdg kolom nama_mhs.
- c. periode_ukmppdg berisi periode ujian UKMP2DG yang diambil dari tabel data_ukmppdg kolom periode.
- d. tanggal_ukmppdg berisi tanggal ujian UKMP2DG yang diambil dari tabel data_ukmppdg kolom tanggal.
- e. status_peserta_ujian yang berisi data hasil transformasi dari tanggal ujian UKMP2DG dengan membandingkan tanggal ujian dengan tanggal ujian pertama. Jika tanggal ujian sama dengan tanggal pertama maka status menjadi First Taker, jika tidak maka Retaker.
- f. status_lulus_sarjana berisi data status lulus mahasiswa PSKG diambil dari tabel mahasiswa_lulus_sarjana kolom status_lulus.
- g. ipk_sarjana berisi data IPK mahasiswa PSKG yang diambil dari tabel mahasiswa_lulus_sarjana kolom ipk.

- h. ipk_profesi berisi data IPK mahasiswa PPDG yang diambil dari tabel mahasiswa_lulus_profesi kolom ipk_pro.
 - i. status_kelulusan berisi data hasil dari transformasi dari kolom kompetensi. Jika kompetensi bernilai KOMPETEN maka rubah menjadi Lulus dan jika TIDAK KOMPETEN menjadi Tidak Lulus
6. Tabel dataset_mahasiswa yang telah dibuat selanjutnya dilakukan validasi untuk memastikan data yang telah dibuat sesuai dengan melakukan *query* seperti pada Gambar 3. 5. Validasi hasil membuat dataset_mahasiswa



Showing rows 0 - 7 (8 total. Query took 0.0085 seconds.)

```
SELECT 'Total dataset_mahasiswa' AS Keterangan, (SELECT COUNT(*) FROM dataset_mahasiswa) jumlah UNION ALL SELECT 'Total data_ukmppdg' AS Keterangan, (SELECT COUNT(*) FROM data_ukmppdg) jumlah UNION ALL SELECT 'Jumlah KOMPETEN' AS Keterangan, (SELECT COUNT(*) FROM data_ukmppdg WHERE kompetensi = 'KOMPETEN') jumlah UNION ALL SELECT 'Jumlah LULUS' AS Keterangan, (SELECT COUNT(*) FROM dataset_mahasiswa WHERE status_kelulusan = 'LULUS') jumlah UNION ALL SELECT 'Jumlah TIDAK KOMPETEN' AS Keterangan, (SELECT COUNT(*) FROM data_ukmppdg WHERE kompetensi = 'TIDAK KOMPETEN') jumlah UNION ALL SELECT 'Jumlah TIDAK LULUS' AS Keterangan, (SELECT COUNT(*) FROM dataset_mahasiswa WHERE status_kelulusan = 'TIDAK LULUS') jumlah UNION ALL SELECT 'Jumlah First Taker' AS Keterangan, (SELECT COUNT(*) FROM dataset_mahasiswa WHERE status_peserta_ujian = 'First Taker') jumlah UNION ALL SELECT 'Jumlah Retaker' AS Keterangan, (SELECT COUNT(*) FROM dataset_mahasiswa WHERE status_peserta_ujian = 'Retaker') jumlah;
```

Extra options

Keterangan	Jumlah
Total dataset_mahasiswa	564
Total data_ukmppdg	564
Jumlah KOMPETEN	421
Jumlah LULUS	421
Jumlah TIDAK KOMPETEN	143
Jumlah TIDAK LULUS	143
Jumlah First Taker	425
Jumlah Retaker	139

Gambar 3. 5. Validasi hasil membuat dataset_mahasiswa

Pada Gambar 3. 5. Validasi hasil membuat dataset_mahasiswa merupakan hasil validasi dataset_mahasiswa. Hasil tersebut menunjukkan bahwa dataset_mahasiswa yang dibuat telah sesuai karena data yang dihasilkan pada dataset_mahasiswa dan data_ukmppdg memiliki jumlah baris yang sama yaitu 564 baris. Pada proses transformasi juga menghasilkan data yang sama antara jumlah baris yang berkategori KOMPETEN sama dengan yang berkategori LULUS dan kategori TIDAK KOMPETEN sama dengan TIDAK LULUS. Kemudian untuk transformasi peserta ujian juga memiliki hasil yang sama dengan jumlah baris dari keseluruhan data_ukmppdg dimana pada kategori First Taker berjumlah 425 dan Retaker 139, maka jika dijumlahkan hasilnya 564 baris.

7. Data pada tabel dataset_mahasiswa yang telah tervalidasi kemudian di *export* dalam format *.csv* untuk digunakan sebagai data untuk penelitian ini.

3.3. Perancangan Sistem Secara Umum

Sistem yang dikembangkan penulis merupakan sistem ramalan ketepatan waktu kelulusan peserta didik berbasis algoritma *Random Forest*, dengan data sampel dari mahasiswa Program Profesi Dokter Gigi. Sistem ini menerima masukan data untuk diproses dan menghasilkan prediksi kelulusan secara otomatis.

3.3.1. Proses Prediksi Kelulusan Tepat Waktu dengan Random Forest



Gambar 3.6. Flowchart Alur Sistem

Berdasarkan Gambar 3.6. *Flowchart* perancangan sistem, Proses diawali dengan upload dataset yang akan digunakan dimasukkan ke dalam sistem berupa data IPK Sarjana dan IPK Profesi. Kemudian dataset dipartisi menjadi dua bagian, yaitu 80% untuk tahap latih dan 20% untuk proses uji, diikuti proses latih dan testing dengan algoritma *Random Forest*. Selanjutnya mengevaluasi model dengan

mencari nilai metrik akurasi, presisi, *recall*, serta *F1-score* guna mengukur performa model. Jika model telah mencapai hasil optimal, maka disimpan untuk prediksi kelulusan mahasiswa selanjutnya; namun bila kinerja belum memadai, pelatihan diulang hingga optimal. Setelah tersimpan, model siap memproses data baru dan menampilkan prediksi mahasiswa berpotensi lulus tepat atau tidak.

3.3.1. Aktor

Pada sistem terdapat 2 (dua) aktor yaitu Admin dan *User* (Dosen dan Mahasiswa) yang dapat melakukan aktivitas dalam sistem.

1. Admin

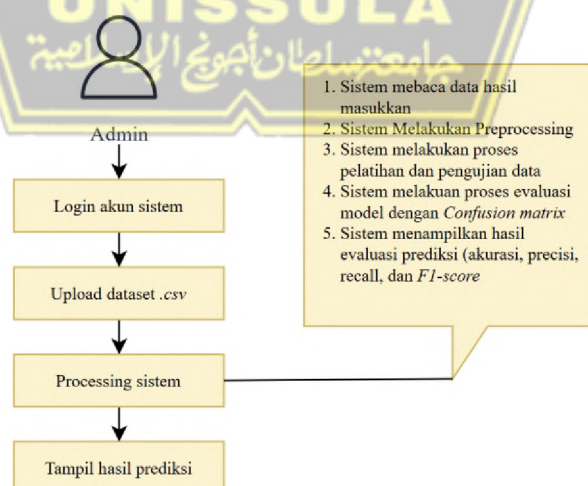
Admin yaitu aktor yang mampu menjalankan berbagai fungsi dalam sistem, sebagaimana terlihat pada aktivitas admin.

2. *User*

User yaitu aktor yang dapat mengoperasikan beberapa fungsi yang terdapat dalam sistem seperti yang terlihat pada aktivitas *user*.

3.3.2. Alur kerja Sistem

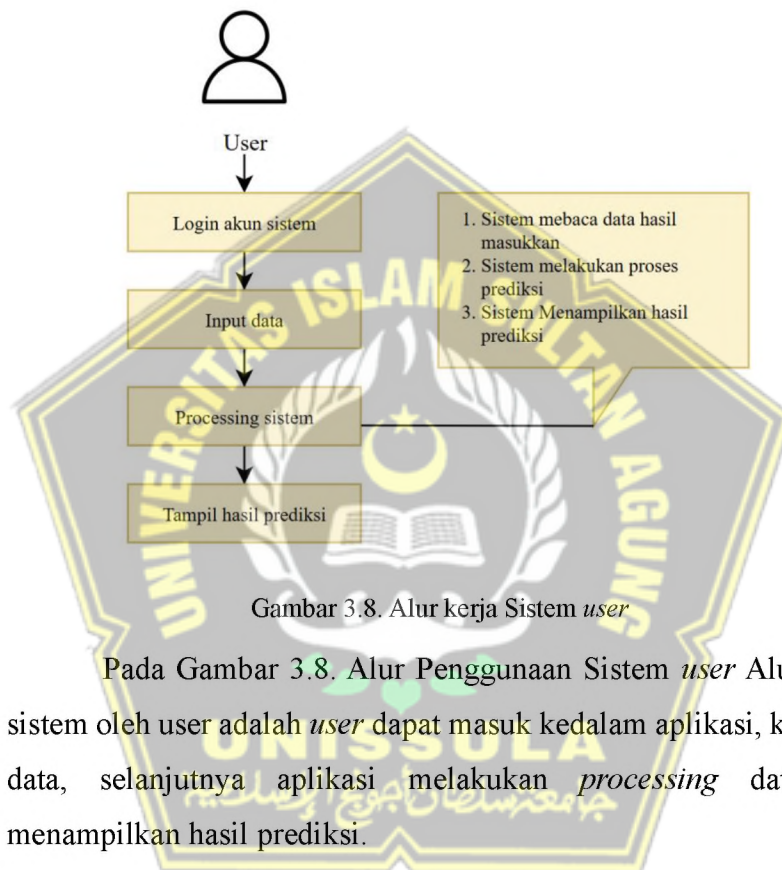
1. Admin



Gambar 3.7. Alur kerja sistem Admin

Pada Gambar 3.7. Alur kerja sistem Admin, pertama melakukan login akun sistem, kemudian admin melakukan upload file dataset berekstensi *.csv* , selanjutnya sistem melakukan *processing* data untuk prediksi kelulusan, kemudian sistem menampilkan hasil prediksi.

2. *User*

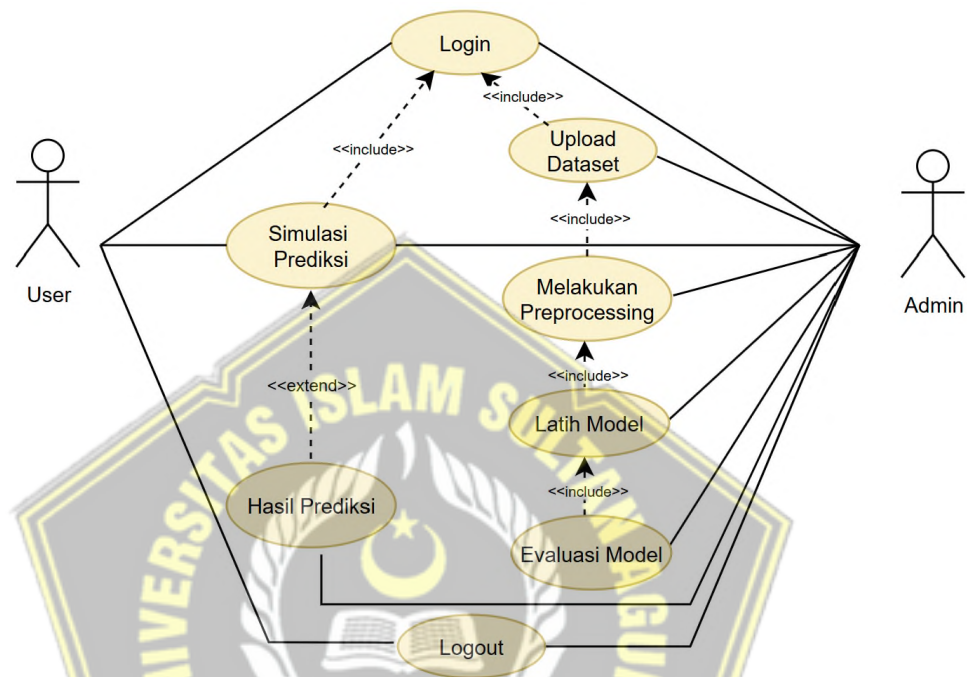


Gambar 3.8. Alur kerja Sistem *user*

Pada Gambar 3.8. Alur Penggunaan Sistem *user* Alur penggunaan sistem oleh *user* adalah *user* dapat masuk kedalam aplikasi, kemudian input data, selanjutnya aplikasi melakukan *processing* data, kemudian menampilkan hasil prediksi.

3.4. Analisa Kebutuhan

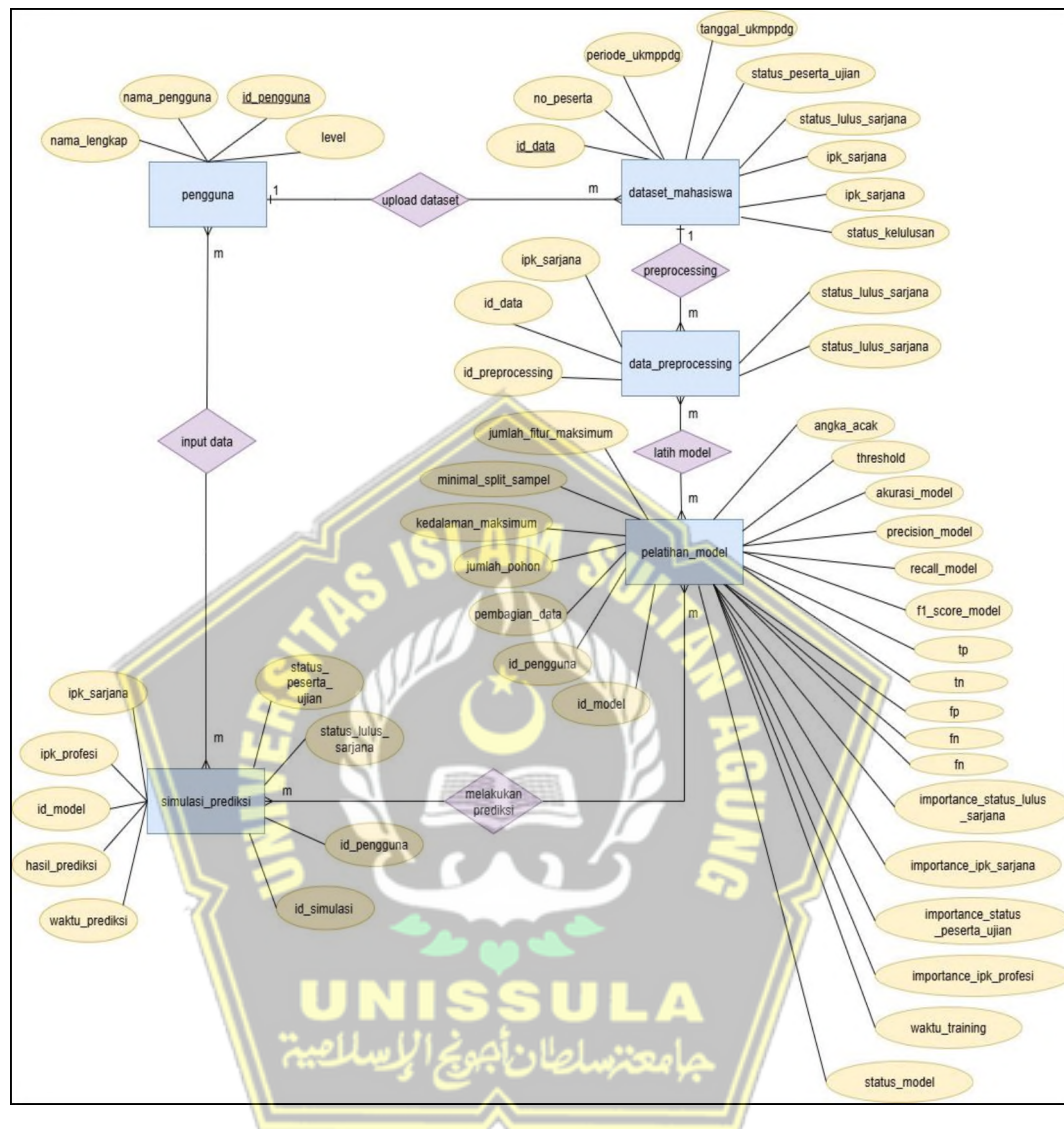
3.4.1. Use Case Diagram



Gambar 3.9. Use Case Diagram

Pada Gambar 3.9. *Use Case Diagram* merupakan *use case* sistem prediksi kelulusan, dalam diagram tersebut terdapat 2 (dua) aktor yang terlibat dalam sistem yaitu *user* yang dikategorikan sebagai mahasiswa atau dosen, admin sebagai aktor sistem untuk mengelola sistem. Aktor *user* dapat melakukan autentikasi masuk, dan keluar dengan akun yang telah terdaftar dalam sistem, melakukan simulasi prediksi dan melihat hasil prediksi. Aktor admin dapat melakukan autentikasi masuk dan keluar, melaksanakan proses upload dataset, *preprocessing* serta melakukan pelatihan model prediksi.

3.4.2. Entity Relation Diagram (ERD)



Gambar 3.10. Entity Relation Diagram (ERD)

Pada Gambar 3.10. Entity Relation Diagram terdapat beberapa entitas yaitu entitas pengguna yang berfungsi menyimpan data pengguna, entitas dataset_mahasiswa merupakan hasil input data mahasiswa yang hanya dapat dilakukan oleh admin, entitas data_preprocessing yaitu entitas yang menyimpan data hasil *preprocessing* yang dilakukan oleh admin dari dataset_mahasiswa, entitas pelatihan_model yaitu entitas yang menyimpan hasil pelatihan model yang dilakukan oleh admin, entitas simulasi_prediksi yaitu entitas hasil prediksi berdasarkan input data yang dapat dilakukan baik *user* maupun admin.

3.4.3. Perancangan Basis Data

Tabel 3.1. pengguna

Field	Type	Keterangan
id_pengguna	int	Primary key
nama_pengguna	varchar(50)	Not Null
kata_sandi	varchar(255)	Not Null
nama_lengkap	varchar(100)	Not Null
Level	enum('admin', 'dosen', 'mahasiswa')	Not Null

Pada Tabel 3.1. pengguna merupakan tabel penyimpanan data pengguna sistem, dilengkapi kolom nama_pengguna dan kata_sandi untuk proses autentikasi masuk aplikasi prediksi kelulusan.

Tabel 3.2. dataset_mahasiswa

Field	Tipe Data	Keterangan
id_data (Primary)	int	Primary key
no_peserta	varchar(50)	Not Null
nama_mahasiswa	varchar(100)	Not Null
periode_UKMP2DG	varchar(20)	Not Null
tanggal_UKMP2DG	date	Not Null
status_peserta_ujian	varchar(20)	Not Null
status_lulus_sarjana	varchar(20)	Not Null
ipk_sarjana	decimal(3,2)	Not Null
ipk_profesi	decimal(3,2)	Not Null
status_kelulusan	varchar(20)	Not Null

Pada Tabel 3.2. Tabel dataset_mahasiswa merupakan tabel yang digunakan untuk menyimpan data hasil upload dataset oleh aktor admin, yang terdiri dari 10 (sepuluh) kolom.

Tabel 3.3. data_preprocessing

Field	Type Data	Keterangan
id_preprocessing	int	Primay key
id_data	int	Foreign key
ipk_sarjana	decimal(3,2)	Not Null
status_lulus_sarjana	tinyint(1)	Not Null
ipk_profesi	decimal(3,2)	Not Null
status_peserta_ujian	tinyint(1)	Not Null
status_kelulusan	tinyint(1)	Not Null

Pada Tabel 3.3 data_preprocessing merupakan tabel yang menyimpan hasil *preprocessing* dari dataset_mahasiswa sebagai data yang akan digunakan dalam pelatihan model.

Tabel 3.4. pelatihan_model

Field	Type	Keterangan
id_model	int	Primary key
id_pengguna	int	Foreign key
pembagian_data	varchar(10)	Not Null
jumlah_pohon	int	Not Null
kedalaman_maksimum	int	Not Null
minimal_split_sampel	int	Not Null
jumlah_fitur_maksimum	int	Not Null
angka_acak	int	Not Null
threshold	float	Not Null
akurasi_model	float	Not Null
precision_model	float	Not Null
recall_model	float	Not Null

f1_score_model	float	Not Null
tp	int	Not Null
tn	int	Not Null
fp	int	Not Null
fn	int	Not Null
importance_status_lulus_sarjana	float	Not Null
importance_ipk_sarjana	float	Not Null
importance_status_peserta_ujian	float	Not Null
importance_ipk_profesi	float	Not Null
waktu_training	datetime	Not Null
status_model	varchar(50)	Not Null

Pada Tabel 3.4. pelatihan_model merupakan tabel yang menyimpan hasil pelatihan model yang dilakukan oleh aktor admin yang terdapat 23 kolom pada pelatihan_model.

Tabel 3.5. simulasi_prediksi

Field	Type	Keterangan
id_simulasi	int	Primary key
id_pengguna	int	Foreign key
id_mahasiswa	int	Foreign key
status_lulus_sarjana	tinyint(1)	Not Null
ipk_sarjana	decimal(3,2)	Not Null
status_peserta_ujian	tinyint(1)	Not Null
ipk_profesi	decimal(3,2)	Not Null
id_model	int	Foreign key
hasil_prediksi	varchar(50)	Not Null
waktu_simulasi	datetime	Yes

Pada Tabel 3.5. Tabel simulai_prediksi digunakan untuk menyimpan riwayat dari hasil prediksi yang dilakukan oleh pengguna.

3.4.4. Data Flow Diagram (DFD)

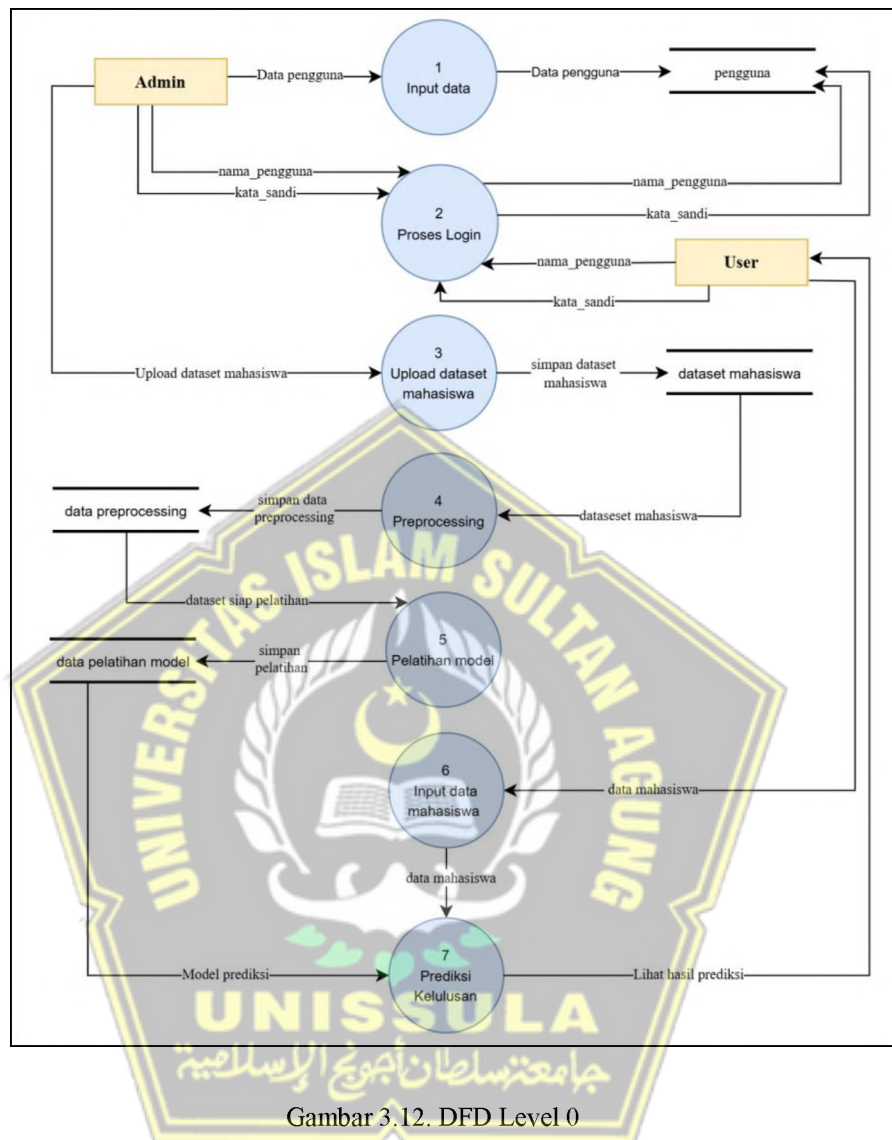
1. Diagram Konteks



Gambar 3.11. Diagram Konteks

Gambar 3.11. Diagram konteks terlihat ada 2 (dua) aktor terdiri dari admin dan *user* yang melakukan aktivitas sistem prediksi kelulusan tepat waktu. Admin *user* melakukan *login* untuk masuk ke dalam sistem, kemudian melakukan input data pengguna dan melakukan pengelolaan data pengguna. Admin melakukan *upload* dataset mahasiswa, menampilkan dataset mahasiswa, mengolah dataset mahasiswa untuk dilakukan pelatihan model, kemudian admin akan dapat melihat hasil evaluasi model yang dilatih, admin dan *user* dapat melakukan prediksi kelulusan tetap waktu dan memeriksa hasil prediksi. Selanjutnya, admin maupun pengguna dapat keluar dari sistem dengan melakukan *logout*.

2. DFD Level 0



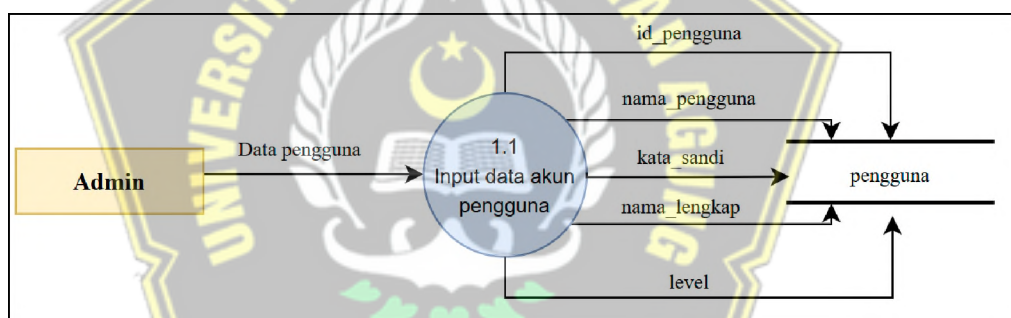
Gambar 3.12. DFD Level 0

Pada Gambar 3.12. DFD level 0 menggambarkan alur diagram DFD level 0 yang terdapat 5 (lima) proses pada sistem prediksi kelulusan tepat waktu yaitu pada admin menjalankan input data pengguna serta proses autentikasi masuk, upload dataset mahasiswa, *preprocessing*, pelatihan model. Pada aktor *user* dapat melakukan input data mahasiswa dan melakukan prediksi kelulusan.

Tahap pertama hak khusus aktor admin dengan melakukan input pengguna yang digunakan sebagai akses masuk sistem oleh user pada proses kedua. Proses selanjutnya admin melakukan proses *upload* dataset mahasiswa berupa *file* berekstensi *.csv*. *File* yang telah diupload admin akan melakukan

proses keempat yaitu *preprocessing* untuk melakukan pembersihan data, menghapus label yang tidak diperlukan, transformasi data kategori ke *biner*, serta melakukan pengecekan data kosong. Data tersebut selanjutnya disimpan ke dalam tabel *preprocessing*. Data *Preprocessing* pada proses kelima admin melakukan pelatihan untuk mendapat model prediksi terbaik kemudian di simpan. Selanjutnya pada proses keenam mahasiswa dapat melakukan input data mahasiswa berupa NIM, Nama, IPK Sarjana, Status Lulus Sarjana, IPK Profesi, dan Status Peserta Ujian UKMP2DG. Kemudian user dapat melakukan proses keenam prediksi kelulusan berdasarkan model prediksi yang telah ditetapkan oleh admin dan user akan dapat melihat hasil prediksi tersebut.

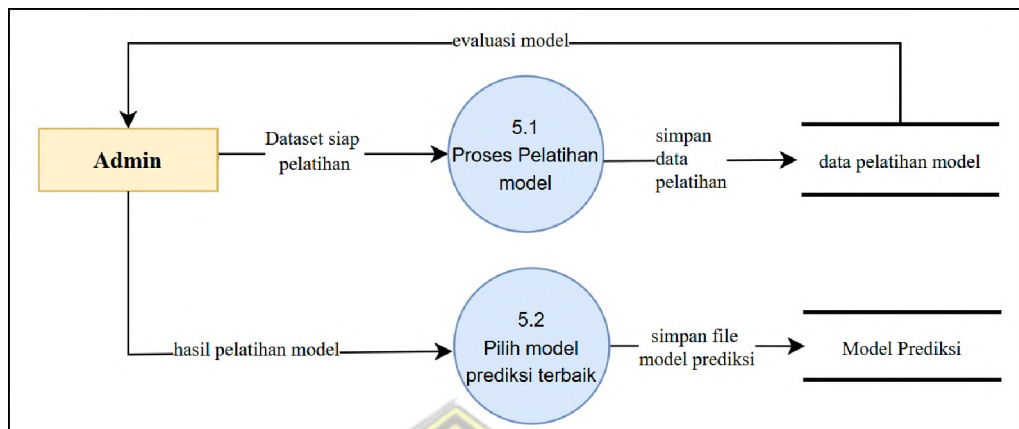
3. DFD Level 1 Proses 1



Gambar 3.13. DFD Level 1 Proses 1

Pada Gambar 3.13. DFD Level 1 Proses 1, aktor admin memasukkan data pengguna yang disimpan dalam basis data pada tabel pengguna, untuk kemudian digunakan dalam proses akses masuk sistem oleh *user*.

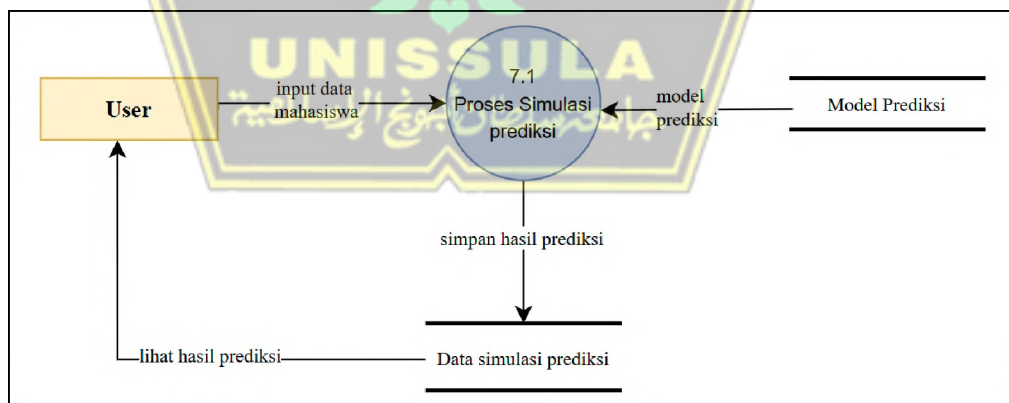
4. DFD Level 1 Proses 5



Gambar 3.14. DFD Level 1 Proses 5

Pada Gambar 3.14. DFD Level 1 Proses 5 merupakan sub proses 5.1 aktor Admin memulai proses melatih model dengan data yang sudah matang dan siap diproses, kemudian hasilnya akan disimpan dalam *database* pada tabel pelatihan model. Aktor admin pada sub proses 5.2 akan melakukan pemilihan model yang terbaik yang kemudian disimpan dalam bentuk *file* berekstensi *.model*.

5. DFD Level 1 Proses 7



Gambar 3.15. DFD Level 1 Proses 7

Pada Gambar 3.15. DFD Level 1 Proses 7, merupakan sub proses 7.1 dimana aktor *user* melakukan proses simulasi prediksi kelulusan berdasarkan input data mahasiswa dengan menggunakan model prediksi yang hasilnya disimpan

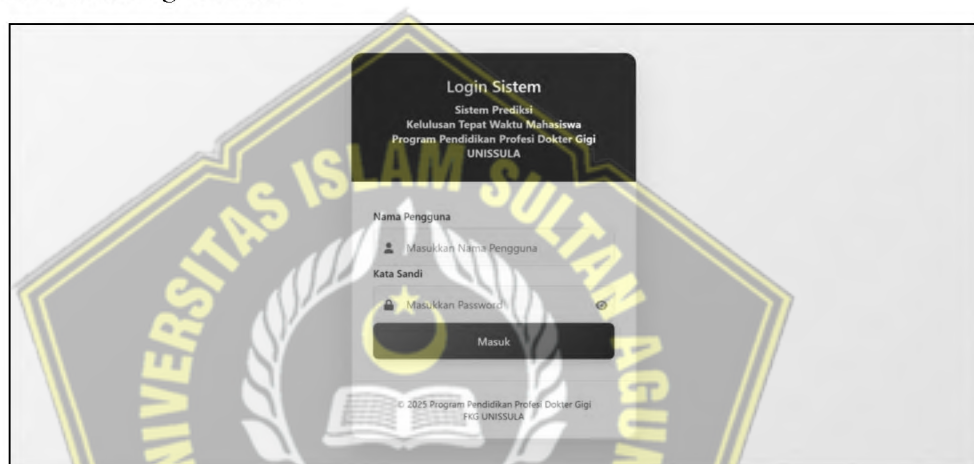
dalam *database* pada tabel simulasi prediksi, kemudian *user* dapat melihat hasil prediksi yang telah dilakukan.

3.5. Perancangan Sistem

3.5.1. Perancangan Desain Sistem

a) Aktor Admin

1. Halaman *Login* Admin



Gambar 3.16. Halaman *login* aktor admin

Pada Gambar 3.16 Halaman login aktor admin, admin melakukan autentikasi login dengan memasukkan username dan password, kemudian bisa masuk kedalam sistem prediksi kelulusan tepat waktu mahasiswa program profesi dokter gigi.

2. Halaman Pengguna

No	Nama Pengguna	Nama Lengkap	Level	Aktif
1	2110002	Nugroho	Mahasiswa	<input checked="" type="checkbox"/> Aktif <input type="button" value="Hapus"/>
2	2110001	Dedi	Mahasiswa	<input checked="" type="checkbox"/> Aktif <input type="button" value="Hapus"/>
3	3100001	Dosen	Dosen	<input checked="" type="checkbox"/> Aktif <input type="button" value="Hapus"/>
4	admin	admin	Admin	<input checked="" type="checkbox"/> Aktif <input type="button" value="Hapus"/>

Gambar 3.17. Halaman data pengguna

Pada Gambar 3.17 Halaman data pengguna, aktor admin dapat melihat data pengguna dan dapat melakukan input pengguna melalui form input pengguna seperti pada Gambar 3. 13 Form input data pengguna.

Tambah Pengguna Baru ×

Nama Pengguna

Nama Lengkap

Kata Sandi

Level

Gambar 3.18. Form input data pengguna

Pada Gambar 3.18. Form input data pengguna, aktor admin melakukan aktivitas input data pengguna dimana data pengguna tersebut

sebagai akses masuk kedalam sistem prediksi kelulusan tepat waktu program Pendidikan profesi.

3. Halaman Dataset

ID	No Peserta	Nama Mahasiswa	Periode	Tanggal	Status Peserta Ujian	IPK Sarjana	Status Lulus Sarjana	IPK Profesi	Status Kelulusan
1	20244102112	Jhon	Periode 4 2024	18 November 2024	Relaker	3,00	Tidak Tepat Waktu	3,27	Lulus
2	20244102213	Shamim	Periode 1 2024	07 February 2024	First Take	3,08	Tepat Waktu	3,27	Lulus
3	20244102214	Mikael	Periode 1 2024	07 February 2024	First Take	3,10	Tepat Waktu	3,36	Tidak Lulus
4	20244102215	Joel M	Periode 2 2024	21 June 2024	First Take	3,10	Tepat Waktu	3,36	Lulus
5	20244102215	Dus Max	Periode 1 2022	18 February 2022	Relaker	3,18	Tepat Waktu	3,45	Lulus

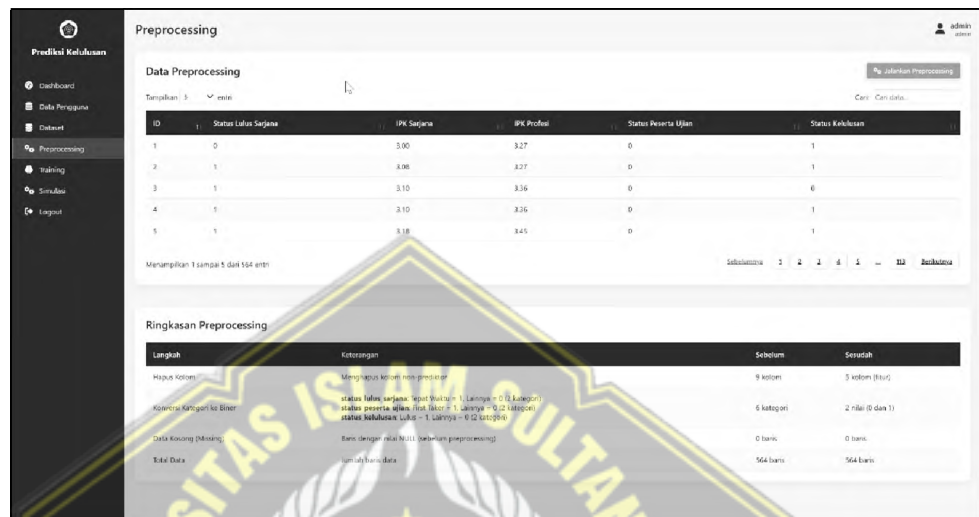
Gambar 3.19. Halaman dataset

Pada Gambar 3.19. Halaman dataset, merupakan antarmuka yang dapat diakses oleh aktor admin untuk melihat dataset dalam bentuk data tabel serta melihat grafik distribusi kelulusan, sebaran ipk, korelasi ipk dan tren kelulusan, serta melakukan aksi upload dataset dengan tombol *upload* data kemudian akan tampil *form* seperti pada Gambar 3.15.

Gambar 3.20. Form Upload dataset

Pada Gambar 3.20. Form Upload dataset, sebagai tempat upload dataset oleh aktor admin dengan memasukkan data berekstensi .csv.

4. Halaman *Preprocessing*



Data Preprocessing

ID	Status Lulus Sarjana	IPK Sarjana	IPK Profesi	Status Peserta Ujian	Status Kelulusan
1	0	3.00	3.27	0	1
2	1	3.08	3.27	0	1
3	1	3.10	3.36	0	0
4	1	3.10	3.36	0	1
5	1	3.18	3.45	0	1

Menampilkan 1 sampai 5 dari 154 entri

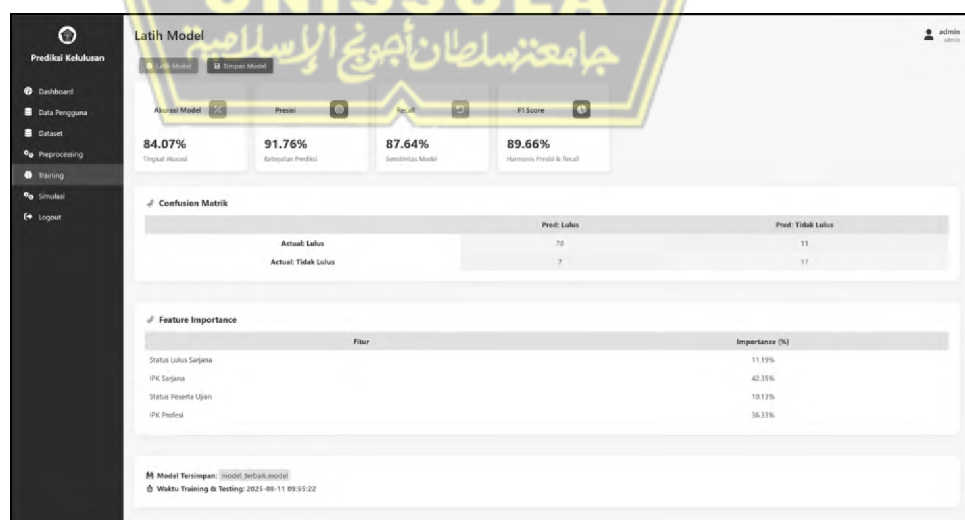
Ringkasan Preprocessing

Langkah	Keterangan	Sebelum	Setelah
Hapus Kolom	Menghapus kolom yang tidak diperlukan	9 kolom	3 kolom (hasil)
Konversi Kategori ke Biner	status lulus sarjana -> Spasi Waktu = 1, Lainnya = 0 (2 kategori) status peserta ujian (0= tidak, 1= lanjut) => 0 (2 kategori) status kelulusan (Lulus = 1, Lainnya = 0 (2 kategori)	5 kategori	2 nilai (0 dan 1)
Data Missing (Missing)	Baris dengan nilai NULL, sebelum preprocessing	0 baris	0 baris
Baki Data	jumlah baris data	564 baris	564 baris

Gambar 3.21. Halaman *Preprocessing*

Pada Gambar 3.21. Halaman *Preprocessing*, aktor admin melaksanakan *preprocessing*, melihat hasil *preprocessing*.

5. Halaman *Training*



Latih Model

Accuracy Model: 84.07% (Tingkat Akurasi)
Precision: 91.76% (Ketepatan Prediksi)
Recall: 87.64% (Sensitivitas Model)
F1 Score: 89.66% (Harmonis Presisi & Recall)

Confusion Matrik

	Pred: Lulus	Pred: Tidak Lulus
Actual: Lulus	79	11
Actual: Tidak Lulus	7	17

Feature Importance

Fitur	Importance (%)
Status Lulus Sarjana	11.19%
IPK Sarjana	42.35%
Status Peserta Ujian	10.12%
IPK Profesi	36.33%

Model Terbilang: model_brtabak.model
🕒 Waktu Training & Testing: 2025-08-11 09:55:22

Gambar 3.22. Halaman *Training*

Pada Gambar 3.22. Halaman Training, aktor admin dalam melakukan training model dan melihat hasil evaluasi model serta menyimpan model prediksi terbaik yang bisa digunakan.

6. Halaman Simulasi Prediksi

No.	NIM	Nama Mahasiswa	Status Lulus Sarjana	IPK Sarjana	IPK Profesi	Status Peserta Ujian	Hasil Prediksi	Waktu Simulasi
1	2110001	Dedi	1	3.19	1	3.34	Lulus	23 Aug 2025 13:54
2	2110002	Nugroho	1	3.40	1	3.20	Lulus	28 Aug 2025 13:26

Gambar 3.23. Halaman Simulasi Prediksi

Pada Gambar 3.23. Halaman Simulasi Prediksi, aktor admin dapat melakukan prediksi kelulusan dengan memasukkan data yang dipilih berdasarkan data pengguna mahasiswa serta memasukkan data pada form input kemudian dapat melihat hasil prediksi yang dilakukan oleh sistem prediksi kelulusan tepat waktu mahasiswa PPDG di FKG UNISSULA.

a) Aktor User

1. Halaman *Login user*

Antarmuka login aktor *user* memiliki kesamaan pada Gambar 3.11. Halaman login aktor admin dan memiliki fungsi yang sama aktor *user* bisa masuk dengan input *username* dan *password*, kemudian bisa masuk kedalam aplikasi prediksi kelulusan tepat waktu mahasiswa.

2. Halaman Simulasi Prediksi *user*

Form Simulasi

NIM: 2110002
Nama: Nugroho
IPK Sarjana: 3.4
IPK Profesi: 3.2
Status Lulus Sarjana: Tepat Waktu
Status Peserta Ujian: First Take

Hasil Simulasi

simulasi "Lulus"

✓ Lulus Tepat Waktu

NIM: 2110002
Nama: Nugroho
IPK Sarjana: 3.4
IPK Profesi: 3.2
Status Lulus Sarjana: Tepat Waktu
Status Peserta Ujian: First Take

Riwayat Simulasi Prediksi

No	NIM	Nama Mahasiswa	Status Lulus Sarjana	IPK Sarjana	IPK Profesi	Status Peserta Ujian	Hasil Prediksi	Waktu Simulasi
1	2110002	Nugroho	1	3.4	3.2	1	Lulus	23 Aug 2025 13:26

Gambar 3.24. Halaman Simulasi Prediksi *User*

Pada Gambar 3.24. Halaman Simulasi Prediksi *User*, aktor *user* dapat melakukan prediksi dan melihat Riwayat prediksi yang telah dilakukan sebelumnya dalam bentuk data tabel.

BAB 4

HASIL DAN ANALISIS PENELITIAN

4.1. Implementasi perangkat lunak

Pada implementasi ini berguna untuk memastikan sistem berjalan optimal. Penulis mengimplementasikan sebagai berikut :

1. Sistem operasi Windows 11 Home Single Language 64-bit
2. Laragon 2025 versi 8.1.0
3. Mysql Versi 8.4.3
4. PHP versi 8.3.16
5. Visual Studio Code versi 1.103.2
6. Browser Google Chrome Versi 139.0.7258.128 (Official Build) (64-bit)

4.2. Implementasi perangkat keras

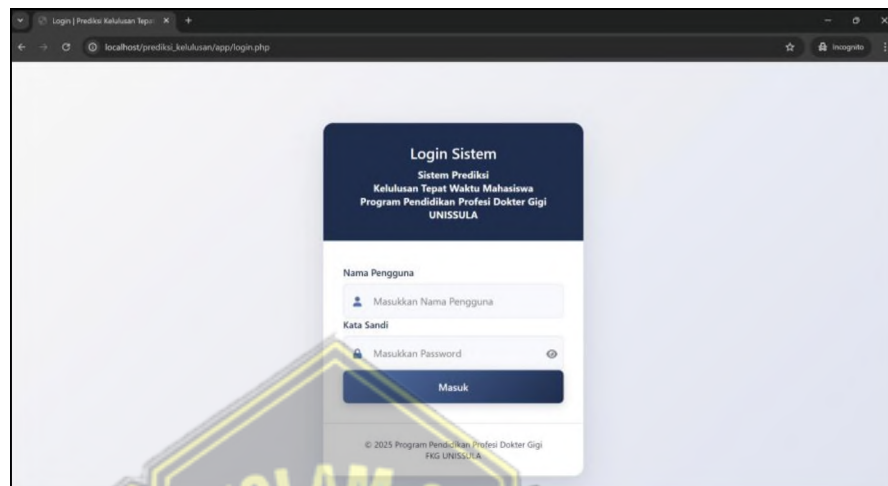
Pada implementasi perangkat keras, untuk memastikan berfungsi optimal sesuai kebutuhan sistem, berikut yang digunakan :

Tabel 4.1. Perangkat Keras

Perangkat keras	Spesifikasi
<i>Processor</i>	13th Gen Intel(R) Core(TM) i9-13900H (20 CPUs)
<i>Memory</i>	DDR4 3200 MHz SDRAM 16 GB
<i>Hardisk</i>	512 GB
<i>Video Graphic Array (VGA)</i>	Intel(R) Iris(R) Xe Graphics
<i>Monitor</i>	14 Inch

4.3. Implementasi *User Interface*

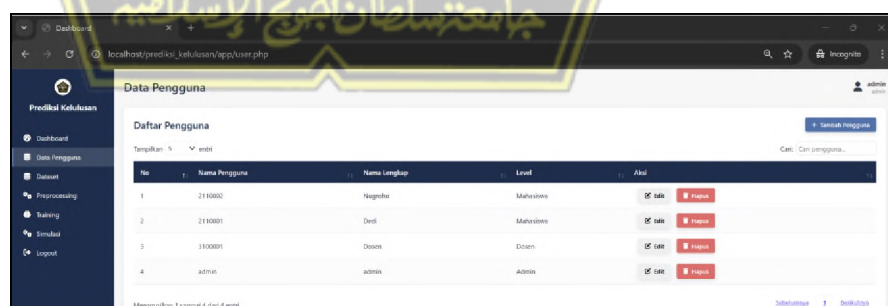
1. Tampilan halaman *User Interface Login* admin dan *user*



Gambar 4.1. *User Interface Login* admin dan user

Pada Gambar 4.1. *User Interface Login* admin dan user merupakan antarmuka halaman login bagi admin dan *user*, berfungsi untuk autentikasi melalui masukkan nama pengguna dan kata sandi guna mengakses aplikasi yang sebelumnya data pengguna sudah dimasukkan oleh admin seperti pada Gambar 4.2. *User Interface data* pengguna.

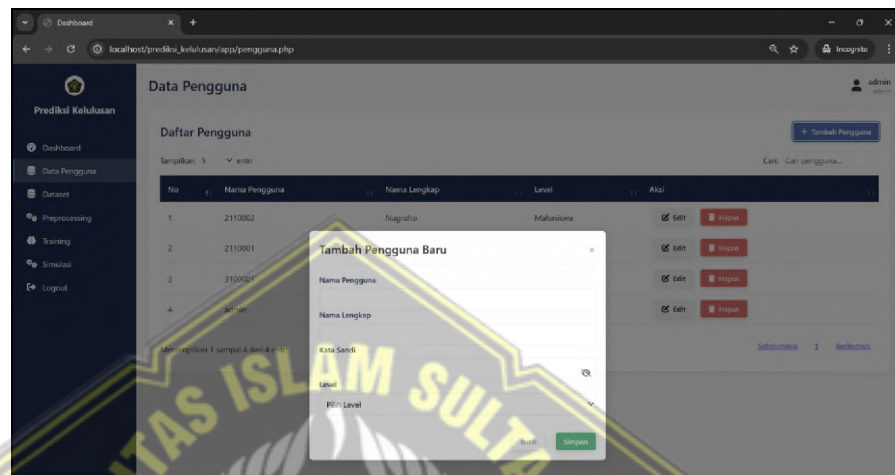
2. *User Interface Data* Pengguna



Gambar 4.2. *User Interface Data* Pengguna

Pada gambar Gambar 4.2. *User Interface Data* Pengguna merupakan antarmuka data pengguna. Pengguna admin dapat melihat informasi semua data dan dapat melakukan aktivitas *Update* dan

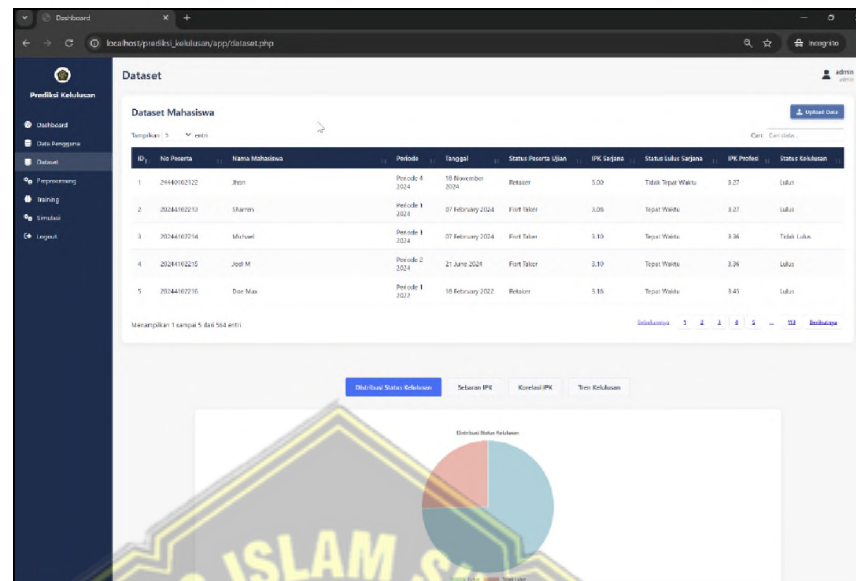
Delete melalui *button edit* dan *hapus*. Admin dapat menambahkan pengguna dengan menekan *button* tambah pengguna untuk memasukan data melalui form input data pengguna seperti pada Gambar 4.3. *User Interface Form Input Pengguna*.



Gambar 4.3. *User Interface Form Input Pengguna*

Pada Gambar 4.3. *User Interface Form Input Pengguna*, admin memiliki hak akses memasukkan nama pengguna, kata sandi, level pengguna, kemudian menyimpan data pengguna yang telah dimasukkan.

3. *User Interface Dataset*



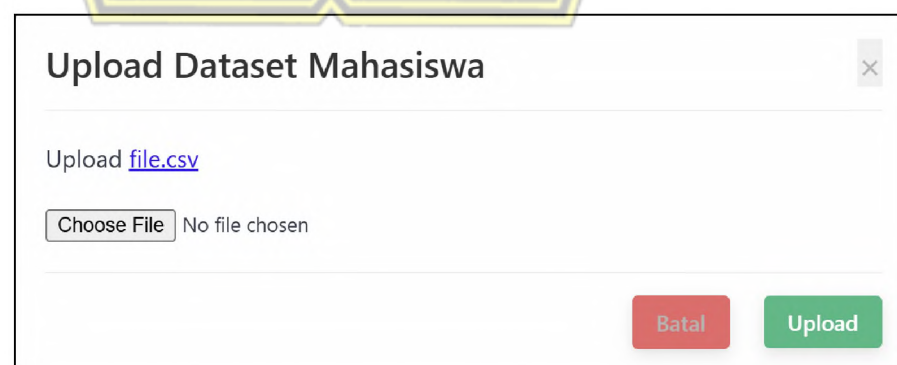
The screenshot shows a web application interface for managing a dataset. The main content area is titled 'Dataset Mahasiswa' and contains a table with the following data:

ID	No Peserta	Nama Mahasiswa	Periode	Tanggal	Status Peserta Ujian	IPK Belajar	Status Lulus Ujian	IPK Profesi	Status Kelulusan
1	20240102122	Jhon	Periode 4 2024	18 Desember 2024	Belajar	3,00	Tidak Tepat Waktu	3,27	Lulus
2	20240102133	Sabrina	Periode 1 2024	07 Februari 2024	Flori Talar	3,20	Tepat Waktu	3,27	Lulus
3	20240102134	Meliani	Periode 1 2024	07 Februari 2024	Flori Talar	3,33	Tepat Waktu	3,36	Tidak Lulus
4	20241102125	Judi M	Periode 2 2024	21 Juni 2024	Flori Talar	3,39	Tepat Waktu	3,26	Lulus
5	20240102135	Dan Nax	Periode 1 2024	18 Februari 2024	Belajar	3,35	Tepat Waktu	3,41	Lulus

Below the table, there is a pie chart titled 'Distribusi Status Kelulusan' with a legend for 'Sebaran IPK', 'Korelasi IPK', and 'Tren Kelulusan'. The chart shows a distribution of data points across these categories.

Gambar 4.4. *User Interface Dataset*

Pada Gambar 4.4. *User Interface Dataset*, merupakan tampilan dari halaman dataset yang digunakan admin untuk *upload file* dataset. Admin dapat melihat dataset, melakukan filter jumlah data yang ditampilkan, melakukan pencarian dataset, serta melihat grafik berdasarkan dataset berupa grafik distribusi status kelulusan, sebaran IPK, korelasi IPK dan tren kelulusan. Dataset diupload melalui form upload seperti pada Gambar 4.5. *User Interface Upload Dataset*.



The screenshot shows a form titled 'Upload Dataset Mahasiswa'. It includes a text input field for the file name, a link to 'Upload file.csv', a 'Choose File' button, and a status indicator 'No file chosen'. At the bottom right, there are two buttons: 'Batal' (Cancel) and 'Upload'.

Gambar 4.5. *User Interface Upload Dataset*

4. *User Interface Preprocessing*

Data Preprocessing

ID	Status Lulus Sarjana	IPK Sarjana	IPK Profesi	Status Peserta Ujian	Status Kelulusan
1	0	3.00	3.27	0	1
2	1	2.08	3.27	0	1
3	1	3.10	3.36	0	0
4	1	3.10	3.36	0	1
5	1	3.18	3.45	0	1

Menampilkan 1 sampai 5 dari 554 entri

Ringkasan Preprocessing

Langkah	Keterangan	Sebelum	Setelah
Hapus Giber	Menghapus kolom non-prediktor	9 kolom	5 kolom (Buru)
Konversi kategori ke Biner	status_lulus_sarjana: 0 (Tidak Lulus) = 0 (2 kategori), 1 (Lulus) = 1 (2 kategori); ipk_sarjana: 0 (Tidak Lulus) = 0 (2 kategori), 1 (Lulus) = 1 (2 kategori); ipk_profesi: 0 (Tidak Lulus) = 0 (2 kategori), 1 (Lulus) = 1 (2 kategori)	6 kategori	2 biner (0 dan 1)
Data Kosong (Missing)	Data dengan nilai NULL telah di-preprocessing	0 baris	0 baris
Total Data	Jumlah baris data	554 baris	554 baris

Gambar 4. 6. *User Interface Preprocessing*

Pada Gambar 4. 5. *User Interface Preprocessing*, merupakan halaman untuk melakukan *preprocessing* yang dapat dilakukan oleh admin untuk melakukan pembersihan data sebelum dilakukan pelatihan model. Dalam halaman tersebut menampilkan data yang sudah dilakukan *preprocessing* serta menampilkan ringkasan dari *preprocessing* yang telah dilakukan oleh admin.

5. *User Interface Training Model*

Latih Model

Akurasi Model: 84.07% (Target Akurasi)

Presisi: 91.76% (Ketepatan Prediksi)

Recall: 87.64% (Sensitivitas Model)

F1 Score: 89.66% (Rata-rata Presisi & Recall)

Confusion Matrik

	Pred: Lulus	Pred: Tidak Lulus
Actual: Lulus	38	11
Actual: Tidak Lulus	7	17

Feature Importance

Fitur	Importance (%)
Status Lulus Sarjana	11.19%
IPK Sarjana	42.35%
Status Peserta Ujian	10.13%
IPK Profesi	36.33%

Waktu Training & Testing: 2025-08-11 09:53:22

Gambar 4. 7. *User Interface Training Model*

Pada halaman yang terlihat pada Gambar 4. 6. *User Interface* Training Model, admin dapat melakukan pelatihan model. Admin dapat melihat hasil pelatihan model, menyimpan kedalam database serta menyimpan model prediksi terbaik dengan ekstensi *.model*.

6. *User Interface* Simulasi Prediksi

No.	NIM	Nama Mahasiswa	Status Lulus Sarjana	IPK Sarjana	IPK Profesi	Status Peserta Ujian	Hasil Prediksi	Waktu Simulasi
1	3202200151	Ende Nugroho	1	3.19	1	3.34	Lulus	31 Aug 2025 22:07
2	2111002	Nugroho	1	3.23	1	3.08	Lulus	30 Aug 2025 14:14

Gambar 4.8. *User Interface* Simulasi Prediksi

Pada Gambar 4.7. *User Interface* Simulasi Prediksi, admin dan user dapat melakukan simulasi prediksi dengan memasukkan data mahasiswa yang akan diprediksi. Untuk admin hanya memilih berdasarkan data pengguna, sedangkan user memasukkan data secara manual kedalam *form* simulasi prediksi kelulusan tepat waktu

4.4. Implementasi Training Model *Random Forest*

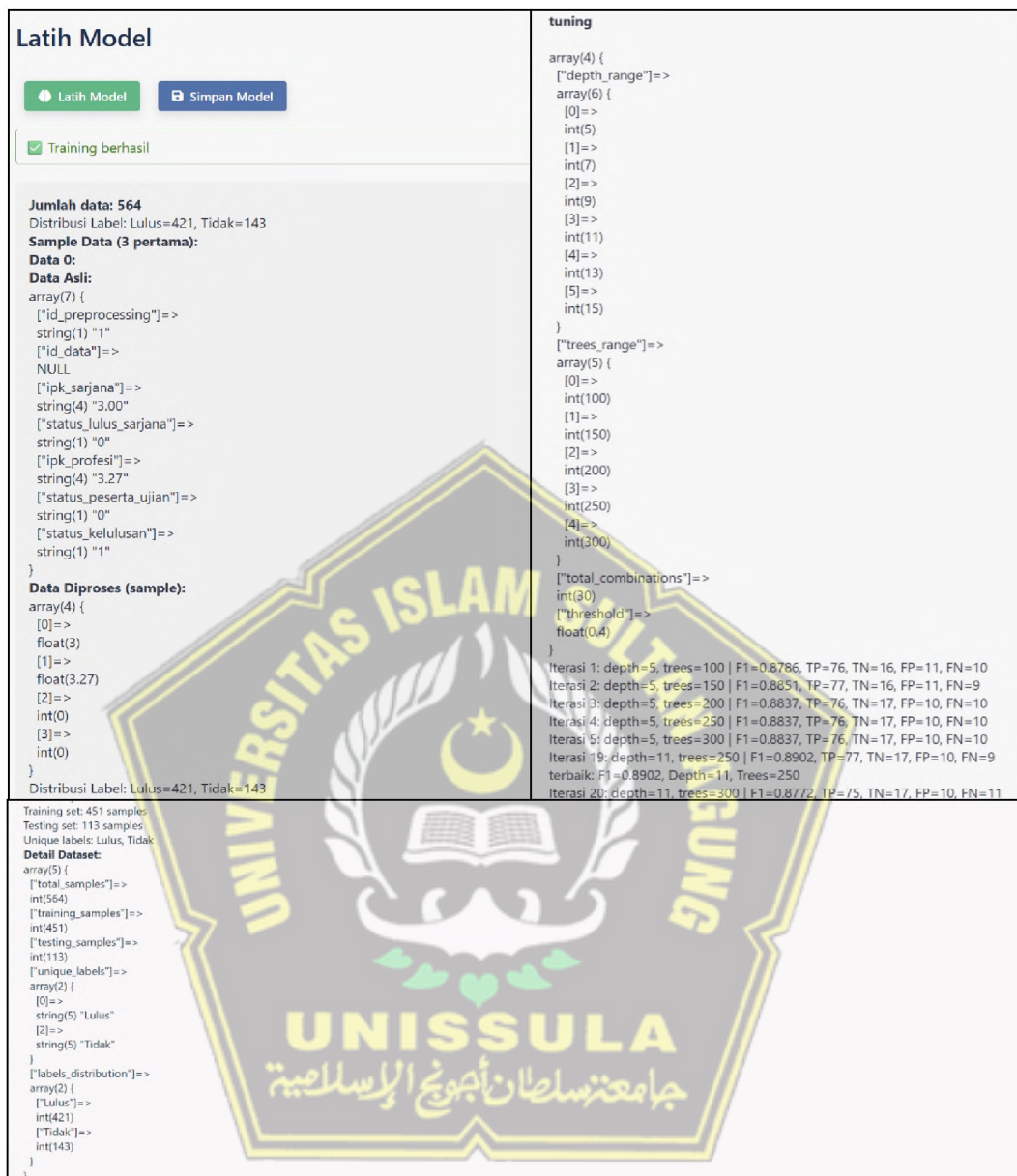
```

130 if (isset($_POST['train'])) {
131     $feature_importances = [];
132     $data = $mysqli->query(query: "SELECT * FROM data_preprocessing");
133     $samples = [];
134     $labels = [];
135     while ($row = $data->fetch_assoc()) {
136         $samples[] = [
137             (float) $row['ipk_sarjana'],
138             (float) $row['ipk_profesi'],
139             (int) $row['status_lulus_sarjana'],
140             (int) $row['status_peserta_ujian']
141         ];
142         $labels[] = $row['status_kelulusan'] == 1 ? 'Lulus' : 'Tidak';
143     }
144
145     if (!empty($samples)) {
146         $startTime = microtime(as_float: true);
147
148         $dataset = new Labeled(samples: $samples, labels: $labels);
149         [$train, $test] = $dataset->randomize()->split(ratio: 0.8);
150
151         $bestF1 = 0;
152         $bestModel = null;
153         $bestParams = [];
154
155         for ($depth = 5; $depth <= 15; $depth += 2) {
156             for ($trees = 100; $trees <= 300; $trees += 50) {
157                 $tree = new ClassificationTree(maxHeight: $depth);
158                 $model = new RandomForest(base: $tree, estimators: $trees, ratio: 0.2, balanced: true);
159                 $model->train(dataset: $train);

```

Gambar 4.9. Potongan Kode Program Training Model

Pada Gambar 4.8. Kode Program Training Model *Random Forest* merupakan potongan kode program proses pelatihan model *Random Forest*. Mulai dengan inisialisasi program untuk menjalankan *train* kemudian mengambil data *preprocessing* dari *database* dengan variabel *samples* yang mengambil fitur IPK Sarjana, IPK Profesi, Status Lulus Sarjana, Status Peserta Ujian serta memberi label status kelulusan menjadi label lulus atau tidak. Kemudian dari *samples* dan label tersebut dibungkus dengan variabel dataset menggunakan kelas *labeled* pada Rubix ML agar bisa dilatih. Dataset tersebut kemudian data diacak, Split data dua komponen, 80% pelatihan dan 20% untuk pengujian. Setelah pemisahan tersebut, proses selanjutnya dilakukan pencarian parameter kedalaman pohon dan jumlah pohon sampai mendapatkan parameter terbaik yang digunakan untuk pembuatan model prediksi kelulusan memanfaatkan kelas *ClassificationTree* serta *RandomForest* pada Rubix ML.



Latih Model

Latih Model Simpan Model

Training berhasil

Jumlah data: 564
Distribusi Label: Lulus=421, Tidak=143

Sample Data (3 pertama):

Data 0:
Data Asli:

```
array(7) {
  ["id_preprocessing"]=>
  string(1) "1"
  ["id_data"]=>
  NULL
  ["ipk_sarjana"]=>
  string(4) "3.00"
  ["status_lulus_sarjana"]=>
  string(1) "0"
  ["ipk_profesi"]=>
  string(4) "3.27"
  ["status_peserta_ujian"]=>
  string(1) "0"
  ["status_kelulusan"]=>
  string(1) "1"
}
```

Data Diproses (sample):

```
array(4) {
  [0]=>
  float(3)
  [1]=>
  float(3.27)
  [2]=>
  int(0)
  [3]=>
  int(0)
}
```

Distribusi Label: Lulus=421, Tidak=143

tuning

```
array(4) {
  ["depth_range"]=>
  array(6) {
    [0]=>
    int(5)
    [1]=>
    int(7)
    [2]=>
    int(9)
    [3]=>
    int(11)
    [4]=>
    int(13)
    [5]=>
    int(15)
  }
  ["trees_range"]=>
  array(5) {
    [0]=>
    int(100)
    [1]=>
    int(150)
    [2]=>
    int(200)
    [3]=>
    int(250)
    [4]=>
    int(300)
  }
  ["total_combinations"]=>
  int(30)
  ["threshold"]=>
  float(0.4)
}
```

Iterasi 1: depth=5, trees=100 | F1=0.6786, TP=76, TN=16, FP=11, FN=10
 Iterasi 2: depth=5, trees=150 | F1=0.8851, TP=77, TN=16, FP=11, FN=9
 Iterasi 3: depth=5, trees=200 | F1=0.8837, TP=76, TN=17, FP=10, FN=10
 Iterasi 4: depth=5, trees=250 | F1=0.8837, TP=76, TN=17, FP=10, FN=10
 Iterasi 5: depth=5, trees=300 | F1=0.8837, TP=76, TN=17, FP=10, FN=10
 Iterasi 19: depth=11, trees=250 | F1=0.8902, TP=77, TN=17, FP=10, FN=9
 terbaik: F1=0.8902, Depth=11, Trees=250
 Iterasi 20: depth=11, trees=300 | F1=0.8772, TP=75, TN=17, FP=10, FN=11

Training set: 451 samples
 Testing set: 113 samples
 Unique labels: Lulus, Tidak

Detail Dataset:

```
array(5) {
  ["total_samples"]=>
  int(564)
  ["training_samples"]=>
  int(451)
  ["testing_samples"]=>
  int(113)
  ["unique_labels"]=>
  array(2) {
    [0]=>
    string(5) "Lulus"
    [2]=>
    string(5) "Tidak"
  }
  ["labels_distribution"]=>
  array(2) {
    ["Lulus"]=>
    int(421)
    ["Tidak"]=>
    int(143)
  }
}
```

Gambar 4. 10. Potongan tampilan implementasi proses training model

Pada Gambar 4. 10. Potongan tampilan implementasi proses training model merupakan tampilan proses training model yang menampilkan data set yang digunakan sejumlah 564 sampel kemudian diambil empat fitur. Selanjutnya Data dipisahkan menjadi dua kelompok utama, 451 untuk pelatihan serta 113 untuk pengujian. Kemudian proses selanjutnya dilakukan proses tuning untuk mencari parameter terbaik dengan mengkombinasikan sebanyak 30 kali. Kemudian setelah selesai akan dipilih hasil yang terbaik.

```

Rubix\ML\Classifiers\RandomForest Object
(
  [base:protected] => Rubix\ML\Classifiers\ClassificationTree Object
  (
    [maxHeight:protected] => 11
    [maxLeafSize:protected] => 3
    [minPurityIncrease:protected] => 1.0E-7
    [root:protected] =>
    [featureCount:protected] =>
    [maxFeatures:protected] =>
    [maxBins:protected] =>
    [classes:protected] => Array
  )
)
[estimators:protected] => 250
[ratio:protected] => 0.2
[balanced:protected] => 1
[trees:protected] => Array
(
  [0] => Rubix\ML\Classifiers\ClassificationTree Object
  (
    [maxHeight:protected] => 11
    [maxLeafSize:protected] => 3
    [minPurityIncrease:protected] => 1.0E-7
    [root:protected] => Rubix\ML\Graph\Nodes\Split Object
    (
      [column:protected] => 1
      [value:protected] => 3.3
      [impurity:protected] => 1.4078474716773
      [n:protected] => 91
      [left:protected] => Rubix\ML\Graph\Nodes\Split Object
      (
        [column:protected] => 2
        [value:protected] => 0
        [impurity:protected] => 1.3229203229203
        [n:protected] => 44
        [left:protected] => Rubix\ML\Graph\Nodes\Split Object
        (
          [column:protected] => 0
          [value:protected] => 3.0456666666667
          [impurity:protected] => 1.268468468468
          [n:protected] => 37
          [left:protected] => Rubix\ML\Graph\Nodes\Split Object
          (
            [column:protected] => 0
            [value:protected] => 2.9866666666667
            [impurity:protected] => 1.1952941176471
            [n:protected] => 25
            [left:protected] => Rubix\ML\Graph\Nodes\Split Object
            (
              [column:protected] => 0
              [value:protected] => 2.862
              [impurity:protected] => 0.25490196078431
              [n:protected] => 17
              [left:protected] => Rubix\ML\Graph\Nodes\Best Object
              (
                [outcome:protected] => Tidak
                [probabilities:protected] => Array
                (
                  [Lulus] => 0.33333333333333
                  [Tidak] => 0.66666666666667
                )
              )
            )
          )
        )
      )
    )
  )
)

```

Gambar 4. 11. Potongan Isi *File* Model Prediksi

Pada Gambar 4.11. Model Prediksi merupakan isi dari *file* model yang telah berhasil dilakukan pada implementasi training model dan telah disimpan dalam bentuk *file* berekstensi *.model* yang dimanfaatkan untuk melakukan prediksi kelulusan tepat waktu.

4.5. Implementasi Perhitungan Evaluasi model

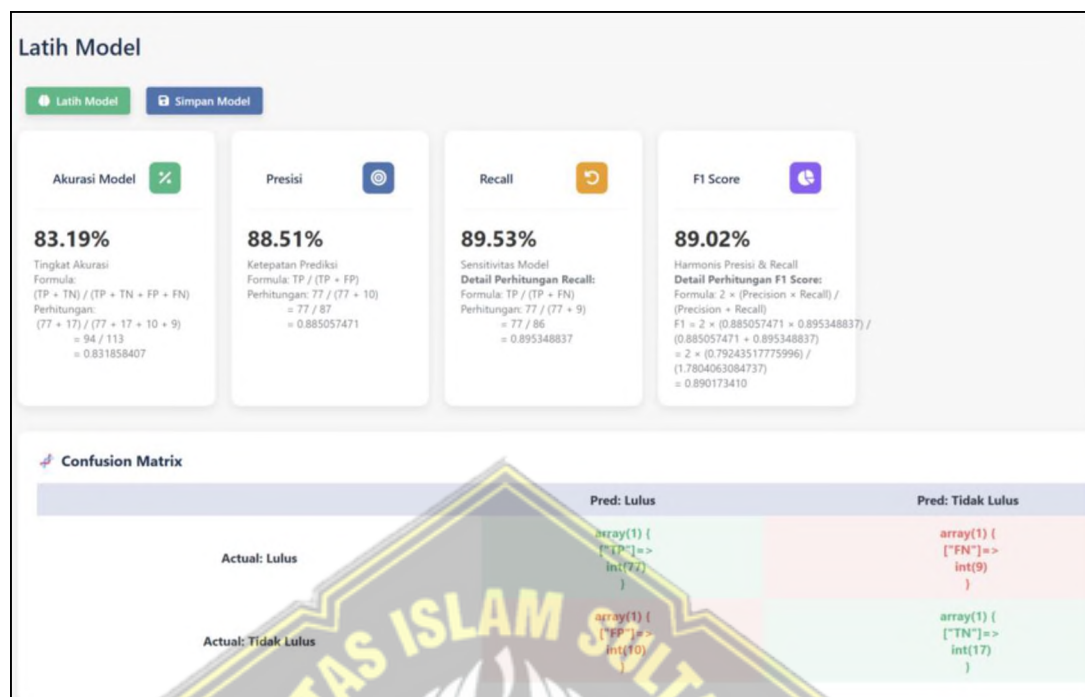
```

170 $tp = $tn = $fp = $fn = 0;
171 foreach ($actual as $i => $true) {
172     $pred = $prediksi[$i];
173     if ($true === 'Lulus' && $pred === 'Lulus') $tp++;
174     elseif ($true === 'Tidak' && $pred === 'Tidak') $tn++;
175     elseif ($true === 'Tidak' && $pred === 'Lulus') $fp++;
176     elseif ($true === 'Lulus' && $pred === 'Tidak') $fn++;
177 }
178
179 $total = count(value: $actual);
180 $prec = $tp / max(value: 1, values: $tp + $fp);
181 $rec = $tp / max(value: 1, values: $tp + $fn);
182 $f1 = 2 * ($prec * $rec) / max(value: 1, values: $prec + $rec);
183
184 if ($f1 > $bestF1) {
185     $bestF1 = $f1;
186     $precision = $prec;
187     $recall = $rec;
188     $f1score = $f1;
189     $akurasi = ($tp + $tn) / max(value: 1, values: $total);
190     $bestModel = $model;
191     $confusion = [
192         'Lulus' => ['Lulus' => $tp, 'Tidak' => $fn],
193         'Tidak' => ['Lulus' => $fp, 'Tidak' => $tn]
194     ];

```

Gambar 4. 12. Potongan Kode Program Implementasi Perhitungan Evaluasi Model

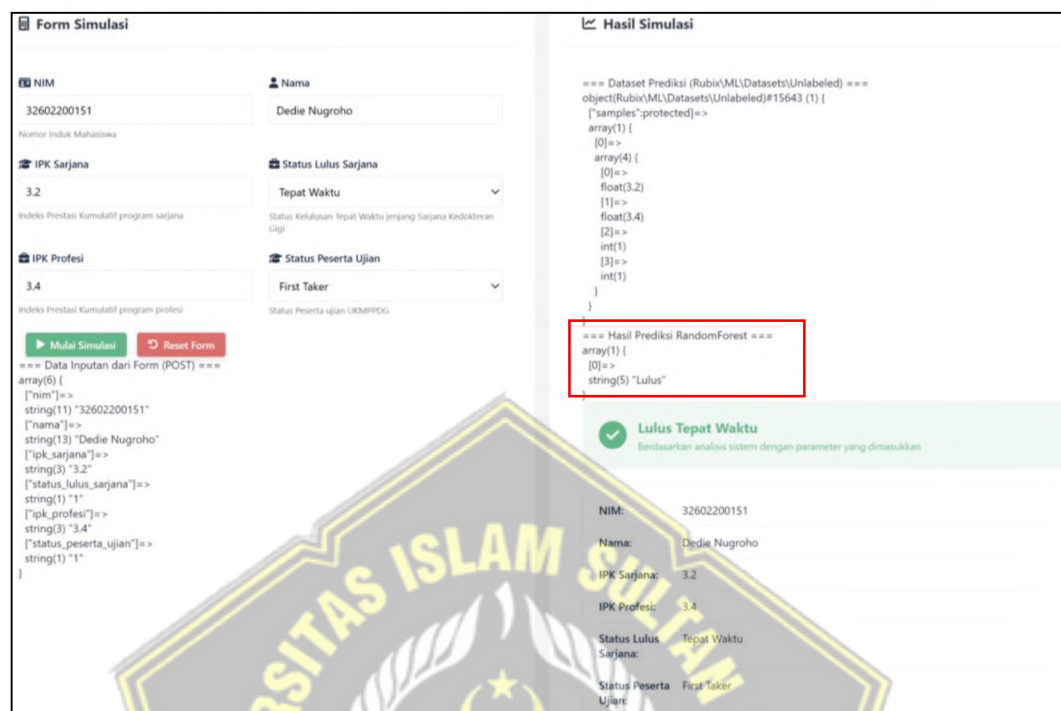
Pada Gambar 4.12. Kode Program Implementasi Perhitungan Evaluasi Model merupakan kode program untuk melakukan perhitungan evaluasi model yang diawali dengan inisialisasi TP, TN, FP dan FN. Kemudian dihitung matriks evaluasi, setelah mendapatkan hasil yang terbaik maka model terbaik disimpan untuk digunakan sistem untuk melakukan prediksi.



Gambar 4.13. Tampilan hasil kode program evaluasi model

Pada Gambar 4.13. Tampilan hasil kode program evaluasi model merupakan tampilan evaluasi model hasil dari kode program pada Gambar 4. 11. Potongan Kode Program Implementasi Perhitungan Evaluasi Model. Dapat terlihat nilai TP, TN, FP, FN yang dihasilkan. kemudian dihitung hasil evaluasi model maka telah diperoleh akurasi 83.19% , presisi 88.51% , recall 89.53% dan F1 Score 89.02% yang diperoleh dari proses sistem yang akan digunakan untuk prediksi kelulusan tepat waktu.

4.6. Implementasi Random Forest Pada Prediksi Kelulusan Tepat Waktu



Form Simulasi

NIM
32602200151
Nomor Induk Mahasiswa

Nama
Dedie Nugroho

IPK Sarjana
3.2
Indeks Prestasi Kumulatif program sarjana

Status Lulus Sarjana
Tepat Waktu
Status Kelulusan Tepat Waktu jenjang Sarjana Kedokteran Gigi

IPK Profesi
3.4
Indeks Prestasi Kumulatif program profesi

Status Peserta Ujian
First Taker
Status Peserta ujian UKMPPOG

Mulai Simulasi **Reset Form**

=== Data Inputan dari Form (POST) ===
array(6) [
 ["nim"]=>
 string(11) "32602200151"
 ["nama"]=>
 string(13) "Dedie Nugroho"
 ["ipk_sarjana"]=>
 string(3) "3.2"
 ["status_lulus_sarjana"]=>
 string(1) "1"
 ["ipk_profesi"]=>
 string(3) "3.4"
 ["status_peserta_ujian"]=>
 string(1) "1"
]

Hasil Simulasi

=== Dataset Prediksi (Rubix\ML\Datasets\Unlabeled) ===
object(Rubix\ML\Datasets\Unlabeled)#15643 (1) [
 ["samples"-protected]>
 array(1) [
 [0]>
 array(4) [
 [0]>
 float(3,2)
 [1]>
 float(3,4)
 [2]>
 int(1)
 [3]>
 int(1)
]
]
]

=== Hasil Prediksi RandomForest ===
array(1) [
 [0]>
 string(5) "Lulus"
]

Lulus Tepat Waktu
Berdasarkan analisis sistem dengan parameter yang dimasukkan

NIM: 32602200151
Nama: Dedie Nugroho
IPK Sarjana: 3.2
IPK Profesi: 3.4
Status Lulus Sarjana: Tepat Waktu
Status Peserta Ujian: First Taker

Gambar 4.14. Prediksi Kelulusan Tepat Waktu

Pada Gambar 4.14. Prediksi Kelulusan Tepat Waktu merupakan tampilan Implementasi Random Forest Pada Prediksi Kelulusan tepat waktu yang diawali dengan memasukkan data NIM, Nama, IPK Sarjana, IPK Profesi dan Status Peserta ujian. Dari data tersebut tersebut kemudian dijalankan dengan model prediksi yang telah dilatih kemudian disimpulkan hasil prediksi berdasarkan pola yang telah dipelajari oleh model.

4.7. Pengujian Sistem

Pada fase pengujian sistem, peneliti menerapkan pendekatan black box testing untuk memverifikasi fitur dan tugas sistem beroperasi sesuai pada rancangan sistem.

1. Pengujian *login* adminTabel 4.2. Pengujian *login* admin

Pengujian halaman login admin (data benar)				
Data Uji	Data Masukkan	Capaian yang diinginkan	Hasil aktual	Hasil pengujian
Nama pengguna : (admin) Kata sandi : (admin)	Nama pengguna : (admin) Kata sandi : (admin)	Menampilkan halaman Simulasi prediksi admin	Berhasil masuk <i>dashboard</i> admin	Berhasil
Pengujian halaman <i>login</i> admin (data salah)				
Data Uji	Data Masukkan	Capaian yang diinginkan	Hasil aktual	Hasil pengujian
Nama pengguna : (admin) Kata sandi : (admin)	Nama pengguna : Admin Kata sandi : 12345	Gagal Menampilkan <i>dashboard</i> admin dan <i>alert Password salah!</i>	Gagal masuk <i>dashboard</i> admin dan menampilkan <i>alert Password salah!</i>	Berhasil

2. Pengujian tambah pengguna admin

Tabel 4.3. Pengujian tambah pengguna admin

Pengujian tambah pengguna admin (data benar)				
Data Uji	Data Masukkan	Capaian yang diinginkan	Hasil aktual	Hasil pengujian
Input nama pengguna : (wajib diisi)	Input nama pengguna: 32602200151	Menampilkan halaman data pengguna dan menampilkan	Berhasil Kembali pada halaman data	Berhasil

Input kata sandi : (wajib diisi)	Input kata sandi : 32602200151	<i>alert data</i> berhasil disimpan	pengguna dan tampil <i>alert data</i> berhasil disimpan	
Input Nama Lengkap : (wajib diisi)	Input Nama Lengkap : Dedie Nugoho			
Pilih level pengguna : (wajib dipilih)	Pilih level pengguna: Mahasiswa			
Pengujian tambah pengguna admin (data salah)				
Data Uji	Data Masukkan	Capaian yang diinginkan	Hasil aktual	Hasil pengujian
Input nama pengguna: (wajib diisi)	Input nama pengguna: (kosong)	Gagal menyimpan data pengguna dan	Gagal menyimpan data pengguna	Berhasil
Input kata sandi : (wajib diisi)	Input kata sandi : 32602200151	menampilkan <i>alert please fill out this field</i>	dan tampil <i>alert please fill out this field</i>	
Input Nama Lengkap : (wajib diisi)	Input Nama Lengkap : Dedie Nugroho			
Pilih level pengguna : (wajib dipilih)	Pilih level pengguna : Mahasiswa			

3. Pengujian *login user*Tabel 4.4. Pengujian *login user*

Pengujian halaman <i>login user</i> (data benar)				
Data Uji	Data Masukkan	Capaian yang diinginkan	Hasil aktual	Hasil pengujian
Nama pengguna : 32602200151 Kata sandi : 32602200151	Nama pengguna : 32602200151 Kata sandi : 32602200151	Menampilkan halaman Simulasi prediksi <i>user</i>	Berhasil masuk halaman Simulasi prediksi <i>user</i>	Berhasil
Pengujian halaman <i>login user</i> (data salah)				
Data Uji	Data Masukkan	Capaian yang diinginkan	Hasil aktual	Hasil pengujian
Nama pengguna : 32602200151 Kata sandi : 32602200151	Nama pengguna : 32602200151 Kata sandi : akundedie	Gagal Menampilkan halaman Simulasi prediksi <i>user</i> dan <i>alert Password</i> salah!	Gagal masuk sistem dan menampilkan <i>alert Password</i> salah!	Berhasil

4. Pengujian *upload* dataset admin

Tabel 4.5. Pengujian upload dataset admin

Pengujian <i>upload</i> dataset admin (data benar)				
Data Uji	Data Masukkan	Capaian yang diinginkan	Hasil aktual	Hasil pengujian
<i>Upload</i> dataset : (jenis file .csv,.xlsx,.xls)	<i>Upload</i> dataset : dataset.csv	Menampilkan halaman dataset admin	Berhasil <i>upload</i> dataset dan menampilkan halaman dataset admin	Berhasil
Pengujian <i>upload</i> dataset admin (data salah)				
Data Uji	Data Masukkan	Capaian yang diinginkan	Hasil aktual	Hasil pengujian
<i>Upload</i> dataset : (jenis file .csv)	<i>Upload</i> dataset : dataset.pdf	Gagal menampilkan halaman dataset admin dan menampilkan <i>alert</i> Gagal File tidak berformat .csv	Gagal menampilkan halaman dataset admin dan berhasil menampilkan <i>alert</i> Gagal <i>File</i> tidak berformat .csv	Berhasil

5. Pengujian prediksi kelulusan tepat waktu admin

Tabel 4.6. Pengujian prediksi kelulusan tepat waktu admin

Pengujian prediksi kelulusan tepat waktu admin (data benar)				
Data Uji	Data Masukkan	Capaian yang diinginkan	Hasil aktual	Hasil pengujian
Pilih Nim mahasiswa : (wajib dipilih)	Pilih Nim mahasiswa : 32602200151	Menampilkan hasil prediksi dan menampilkan data hasil prediksi dalam tabel	Berhasil menampilkan hasil prediksi dan menampilkan hasil prediksi dalam tabel	Berhasil
Input Nama Mahasiswa : (otomatis menyesuaikan nim yang dipilih)	Input Nama Mahasiswa : Dedie Nugroho	Menampilkan hasil prediksi dalam tabel	menampilkan hasil prediksi dalam tabel	
Input IPK Sarjana : (wajib diisi dengan minimal IPK lebih besar atau sama dengan 2.75)	Input IPK Sarjana : 3.00			
Pilih status lulus sarjana: (wajib memilih salah satu)	Pilih status lulus sarjana: Tepat Waktu			
Input IPK Profesi : (wajib diisi dengan minimal IPK lebih	Input IPK Profesi : 3.25			

besar atau sama dengan 3.00)				
Pilih status ujian : (wajib memilih salah satu)	Pilih status ujian : First Taker			
Pengujian <i>upload</i> dataset admin (data salah)				
Data Uji	Data Masukkan	Capaian yang diinginkan	Hasil aktual	Hasil pengujian
Nim Input IPK Sarjana : (wajib diisi dengan minimal IPK lebih besar atau sama dengan 2.75) Pilih status lulus sarjana: (wajib memilih salah satu)	Input IPK Sarjana : 2.00 Pilih status lulus sarjana: Tepat Waktu	Gagal Menampilkan hasil prediksi dan menampilkan peringatan <i>value must be greater than or equal to</i> 2.75	Gagal menampilkan hasil prediksi dan berhasil menampilkan peringatan <i>value must be greater than or equal to</i> 2.75	Berhasil
Input IPK Profesi : (wajib diisi dengan minimal IPK lebih besar atau sama dengan 3.00) Pilih status ujian :	Input IPK Profesi : 3.25 Pilih status ujian :			

(wajib memilih salah satu)	First Taker			
----------------------------	-------------	--	--	--

6. Pengujian prediksi kelulusan tepat waktu *user*

Tabel 4.7. Pengujian prediksi kelulusan tepat waktu *user*

Pengujian prediksi kelulusan tepat waktu <i>user</i> (data benar)				
Data Uji	Data Masukkan	Capaian yang diinginkan	Hasil aktual	Hasil pengujian
Input IPK Sarjana : (wajib diisi dengan minimal IPK lebih besar atau sama dengan 2.75)	Input IPK Sarjana : 3.00	Menampilkan hasil prediksi dan menampilkan data hasil prediksi dalam tabel	Berhasil menampilkan hasil prediksi dan menampilkan hasil prediksi dalam tabel	Berhasil
Pilih status lulus sarjana: (wajib memilih salah satu)	Pilih status lulus sarjana: Tepat Waktu			
Input IPK Profesi : (wajib diisi dengan minimal IPK lebih besar atau sama dengan 3.00)	Input IPK Profesi : 3.25			
Pilih status ujian : (wajib memilih salah satu)	Pilih status ujian : First Taker			

Pengujian prediksi kelulusan tepat waktu <i>user</i> (data salah)				
Data Uji	Data Masukkan	Capaian yang diinginkan	Hasil aktual	Hasil pengujian
Input IPK Sarjana : (wajib diisi dengan minimal IPK lebih besar atau sama dengan 2.75)	Input IPK Sarjana : 3.00	Gagal Menampilkan hasil prediksi dan menampilkan peringatan	Gagal menampilkan hasil prediksi dan berhasil menampilkan peringatan	Berhasil
Pilih status lulus sarjana: (wajib memilih salah satu)	Pilih status lulus sarjana: Tepat Waktu	<i>Please select an item in the list</i>	<i>Please select an item in the list</i>	
Input IPK Profesi : (wajib diisi dengan minimal IPK lebih besar atau sama dengan 3.00)	Input IPK Profesi : 3.25			
Pilih status ujian : (wajib memilih salah satu)	Pilih status ujian : Tidak dipilih			

BAB 5

KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan analisis komprehensif yang telah dipaparkan pada Bab 1 hingga Bab 4, penulis merumuskan kesimpulan pada Bab 5 sebagai berikut:

1. Pada penelitian yang telah dilakukan menggunakan empat variabel data akademik, diperoleh kontribusi masing-masing variabel yaitu IPK Sarjana 41.26%, IPK Profesi 36.53%, Status Lulus Sarjana 12.47% dan Status Peserta Ujian 9.74%. Dari hasil tersebut dapat diketahui IPK Sarjana merupakan faktor yang paling berpengaruh terhadap prediksi kelulusan mahasiswa. Hal ini menunjukkan prestasi akademik mahasiswa PSKG menentukan kelulusan tepat waktu mahasiswa PPDG dalam ujian UKMP2DG.
2. Model prediksi dengan algoritma *Random Forest* menunjukkan performa yang baik dengan tingkat akurasi 83.19%, presisi 88.51% , recall 89.53% dan F1 Score 89.02% . Hasil ini menunjukkan model dapat mengidentifikasi mahasiswa beresiko tidak lulus sejak dini.
3. Sistem prediksi berbasis web berhasil dirancang dan diimplementasikan dengan model algoritma *Random Forest* yang dapat digunakan di PPDG FKG UNISSULA sebagai alat untuk mengidentifikasi mahasiswa PPDG yang beresiko tidak lulus sehingga dapat diambil keputusan dan melakukan tindakan khusus kepada mahasiswa yang teridentifikasi.

5.2. Saran

Meskipun penelitian berhasil membangun sistem serta model prediksi, terdapat beberapa keterbatasan yang masih perlu diatasi. Oleh karena itu, penulis mengusulkan sejumlah rekomendasi untuk studi lanjutan guna meningkatkan performa dan keandalan sistem secara signifikan.

1. Penelitian ini masih terbatas hanya menggunakan data akademik, sehingga model belum mengetahui faktor-faktor lain diluar akademik yang mempengaruhi kelulusan mahasiswa yang mengakibatkan 16,81% masih ada kesalahan dalam prediksi. Oleh karena itu perlu menambahkan data non akademik seperti kebiasaan belajar mahasiswa, kondisi ekonomi mahasiswa atau keaktifan mahasiswa dalam organisasi di kampus. Penambahan variabel tersebut diharapkan dapat meningkatkan akurasi model serta dapat mengetahui variabel-variabel lain yang berperan dalam menentukan kelulusan mahasiswa PPDG.
2. Model prediksi yang digunakan masih menggunakan satu model, sehingga belum bisa dipastikan bahwa algoritma *Random Forest* sebagai algoritma terbaik dalam prediksi kelulusan mahasiswa. Penelitian selanjutnya dapat melakukan perbandingan algoritma lain seperti *XGBoost*, *Support Vector Machine (SVM)*, *Neural Network*, atau *Gradient Boosting* agar dapat mengetahui performa keseluruhan algoritma-algoritma tersebut, sehingga dapat ditentukan algoritma yang paling unggul dalam memprediksi kelulusan mahasiswa PPDG.
3. Dataset yang digunakan masih relatif terbatas, sehingga model beresiko *overfitting*. Oleh sebab itu penelitian selanjutnya perlu memperluas jumlah data dengan menambahkan data dari periode-periode UKMP2DG sebelumnya, agar model mampu melakukan prediksi kelulusan mahasiswa lebih akurat.

DAFTAR PUSTAKA

- Amna, A., Wahyuddin, S., Sudipa, I.G.I., Putra, T.A.E., Wahidin, A.J., Syukrilla, W.A., Wardhani, A.K., Heryana, N., Indriyani, T. & Santoso, L.W. 2023. *Data mining*.
- Andrianof, H., Gusman, A.P. & Putra, O.A. 2025. Jurnal Sains Informatika Terapan (JSIT). *Implementasi Algoritma Random Forest untuk Prediksi Kelulusan Mahasiswa Berdasarkan Data Akademik: Studi Kasus di Perguruan Tinggi Indonesia*, 4(1): 24–28.
- Balabied, S.A.A. & Eid, H.F. 2023. Utilizing random forest algorithm for early detection of academic underperformance in open learning environments. *PeerJ Computer Science*, 9: 1–16.
- BPS Indonesia, S.I. 2025. Catalog : 1101001. *Statistik Indonesia 2025*, 1101001: 790. Tersedia di <https://www.bps.go.id/id/publication/2025/02/28/8cfe1a589ad3693396d3db9f/statistik-indonesia-2025.html>.
- Diantika, S., Nalatissifa, H., Maulidah, N., Supriyadi, R. & Fauzi, A. 2024. Penerapan Teknik Random Oversampling Untuk Memprediksi Ketepatan Waktu Lulus Menggunakan Algoritma Random Forest. *Computer Science (CO-SCIENCE)*, 4(1): 11–18.
- Dutt, S.C.A.K.D.S. 2018. *Machine Learning*. PEARSON INDIA.
- Ermatita, H. 2025. Prediksi Kelulusan Mahasiswa Menggunakan Metode K-Means dan Random Forest Fakultas Ilmu Komputer Universitas Sriwijaya , Indonesia Sales Prediction of Student Graduation using K-Means and Random Forest Methods. 5(1): 209–214.
- Ghosh, S.K. & Janan, F. 2021. Prediction of student’s performance using random forest classifier. *Proceedings of the International Conference on Industrial Engineering and Operations Management*, 7089–7100.
- Habibie Sukarna, R., Holilah, Damyati, F. & Hilman, M. 2024. ANALISIS PREDIKSI KELULUSAN MAHASISWA TEPAT WAKTU

MENGGUNAKAN METODE ALGORITMA MACHINE LEARNING
DAN FEATURE SELECTION. 8(2): 232–241.

- Indra, I., Qisthiano, M.R., Syahril, M. & M. Jakak, P. 2023. Optimasi Prediksi Kelulusan Tepat Waktu: Studi Perbandingan Algoritma Random Forest dan Algoritma K-NN Berbasis PSO. *Jurnal Pengembangan Sistem Informasi dan Informatika*, 4(4): 26–35.
- Kemendikbudristek RI 2023. Peraturan Menteri Pendidikan, Kebudayaan, Riset, Dan Teknologi Republik Indonesia Nomor 53 Tahun 2023 Tentang Penjaminan Mutu Pendidikan Tinggi. *Kementerian Pendidikan, Kebudayaan, Riset, Dan Teknologi Republik Indonesia*, (638): 1–45. Tersedia di <https://peraturan.bpk.go.id/Details/265158/permendikbudristek-no-53-tahun-2023>.
- Marwa Sulehu, W. & First Wanita, M. 2025. Optimasi Prediksi Kelulusan Mahasiswa Menggunakan Random Forest untuk Meningkatkan Tingkat Retensi. *Jurnal Minfo Polgan*, 13: 2364–2374. Tersedia di <https://doi.org/10.33395/jmp.v13i2.14472%0Ae-ISSN>.
- Nadia Tarmizi, S. 2022. *Penuhi Kebutuhan Dokter Gigi, Kemenkes Ajak FKG UI Sebar Lulusannya ke 3285 Puskesmas*. Kepala Biro Komunikasi dan Pelayanan Publik. Tersedia di <https://sehatnegeriku.kemkes.go.id/baca/rilis-media/20221208/4141963/penuhi-kebutuhan-dokter-gigi-kemenkes-ajak-fkg-ui-sebar-lulusannya-ke-3285-puskesmas/> [Accessed 19 Mei 2025].
- Nurfadilla, Z. & Faisal 2022. Implementasi Data Mining Untuk Memprediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Random Forest. *Journal of Artificial Intelligence & Data Science*, 01(1): 127–135.
- Nurhalizah, R.S., Ardianto, R. & Purwono, P. 2024. Analisis Supervised dan Unsupervised Learning pada Machine Learning: Systematic Literature Review. *Jurnal Ilmu Komputer dan Informatika*, 4(1): 61–72.
- PDGI 2025. *Statistik*. Persatuan Dokter Gigi Indonesia (PDGI). Tersedia di

<https://pdgi.or.id/halaman/statistik> [Accessed 19 Mei 2025].

Program Profesi Dokter Gigi 2024. Buku Panduan Akademik Profesi Dokter Gigi.

Putra, M. & Harahap, E. 2024. Machine Learning pada Prediksi Kelulusan Mahasiswa Menggunakan Algoritma Random Forest. 127–136.

Riskiyono, F. & Mahdiana, D. 2024. Implementation of Random Forest Algorithm for Graduation Prediction. *Sinkron*, 8(3): 1662–1670.

Ubaidilah, R.M. 2023. Prediksi Kelulusan Mahasiswa Berdasarkan Data Kunjung dan Peminjaman Buku menggunakan Rapid Miner dengan Metode C.45 dan Random Forest. *International Research on Big-Data and Computer Technology: I-Robot*, 7(2): 14–20.

