

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perguruan tinggi adalah lembaga yang menyediakan layanan pembelajaran pada tingkat yang lebih tinggi, yang merupakan tahap terakhir dalam pendidikan formal. Umumnya, perguruan tinggi berbentuk universitas, akademi, institusi, atau politeknik. Terdapat beberapa jenis pendidikan tinggi, yaitu vokasi, akademik, dan profesional. Berdasarkan jenjangnya, pendidikan tinggi menawarkan program diploma, sarjana, magister, spesialis, dan doktor. Perguruan tinggi diharapkan mampu memberikan pendidikan berkualitas kepada mahasiswanya sehingga menghasilkan lulusan yang kompeten di bidangnya. Di Indonesia, kualitas sebuah institusi pendidikan tinggi dinilai berdasarkan akreditasi yang diterbitkan oleh Badan Akreditasi Nasional Perguruan Tinggi (BAN-PT) [1].

Universitas Harkat Negeri sebagai institusi perguruan tinggi swasta memiliki 21 program studi, yaitu D-4 Akuntansi Sektor Publik, D-4 Teknik Informatika, D-3 Akuntansi, D-3 Desain Komunikasi Visual, D-3 Farmasi, D-4 Kebidanan, D-3 Keperawatan, D-3 Perhotelan, D-3 Teknik Elektronika, D-3 Teknik Komputer, D-3 Teknik Mesin, S1 Akuntansi, S1 Hukum, S1 Pend. Guru SD, S1 Psikologi, S1 Sains Data, S1 Ilmu Komunikasi, S1 Teknik Informatika, S1 Teknik Mesin, S1 Manajemen dan S1 Sistem Informasi yang menetapkan standar kelulusan dalam waktu 3-4 tahun (6-8 semester). Namun, masih terdapat sejumlah mahasiswa yang mengalami keterlambatan dalam menyelesaikan studinya. Kondisi ini berdampak

pada penurunan nilai akreditasi universitas karena jumlah lulusan tepat waktu menjadi salah satu parameter penilaian penting oleh lembaga akreditasi [2].

Keterlambatan kelulusan mahasiswa dapat disebabkan oleh berbagai faktor, baik akademik maupun non-akademik. Faktor-faktor tersebut meliputi indeks prestasi semester (IPS), indeks prestasi kumulatif (IPK), masa studi, status bekerja, keikutsertaan organisasi, kegagalan mata kuliah, dan jumlah cuti akademik [3]. Tingkat kelulusan tepat waktu merupakan salah satu indikator utama dalam penilaian akreditasi universitas. Oleh karena itu, dibutuhkan suatu sistem yang mampu memprediksi kelulusan mahasiswa dengan memanfaatkan data profil serta data akademik.

Berdasarkan data yang diperoleh dari Badan Administrasi Akademik (BAA), terdapat sejumlah mahasiswa Program Studi Sarjana Terapan Teknik Informatika yang tidak berhasil menyelesaikan studi dalam waktu 4 tahun. Dari data tersebut, diketahui bahwa beberapa mahasiswa mengalami keterlambatan studi karena faktor akademik seperti rendahnya nilai IPS pada semester awal, serta faktor non-akademik seperti status cuti, dan lama masa studi. Sebagai contoh, mahasiswa yang memiliki status cuti menunjukkan kecenderungan menyelesaikan studi lebih dari 4 tahun. Ini menjadi indikator bahwa variabel non-akademik memiliki peran penting dalam mempengaruhi kelulusan tepat waktu. Data ini juga menunjukkan bahwa terdapat variasi signifikan pada nilai IPS antar semester, dengan beberapa mahasiswa mengalami penurunan nilai dari semester ke semester.

Salah satu pendekatan yang dapat digunakan adalah metode klasifikasi dalam data mining, yang berfungsi untuk mengelompokkan data ke dalam kategori

tertentu berdasarkan data yang telah diberi label (*supervised learning*), sehingga dapat menghasilkan model prediktif yang akurat [4].

Data mining adalah proses mengolah data besar dengan menggunakan metode *statistic*, *machine learning*, dan kecerdasan buatan untuk menemukan pola dan pengetahuan bermanfaat [5]. Salah satu teknikya yaitu klasifikasi, digunakan untuk memodelkan kelas data dan memprediksi kecenderungan, seperti pola kelulusan mahasiswa agar prediksi kelulusan dapat diketahui lebih awal dan mendukung pengambilan keputusan [6].

Sistem prediksi dihitung menggunakan metode algoritma *Decision Tree*. Sistem prediksi menggunakan metode *Decision tree* merupakan sebuah prediksi atau ramalan yang dihitung menggunakan atribut yang dimasukkan kedalam rumus tertentu dan menghasilkan sebuah pohon keputusan yang ditemukan. Metode ini menguraikan faktor-faktor kemungkinan, serta hasil akhir yang ditunjukkan pohon keputusan tersebut [7].

Berbagai penelitian sebelumnya telah menerapkan algoritma *Decision Tree* untuk memprediksi kelulusan mahasiswa dan menunjukkan hasil akurasi yang cukup tinggi, namun sebagian besar belum mengembangkan sistem prediksi yang terintegrasi dan aplikatif. Umumnya, penelitian hanya membandingkan satu atau dua algoritma klasifikasi tanpa eksplorasi metode yang lebih beragam atau optimasi parameter, serta masih berfokus pada variabel akademik semata. Padahal, faktor non-akademik seperti status cuti, keikutsertaan organisasi, status bekerja, dan latar belakang sosial ekonomi juga berpotensi mempengaruhi kelulusan, namun belum dimanfaatkan secara optimal. Selain itu, minimnya pengembangan sistem prediksi

berbasis *web* yang dapat digunakan langsung oleh institusi sebagai alat bantu pengambilan keputusan strategis menjadi kekurangan umum dari penelitian terdahulu. Oleh karena itu, dibutuhkan penelitian lanjutan yang tidak hanya membangun model prediksi yang akurat, tetapi juga merancang sistem prediktif berbasis *website* yang responsif, terintegrasi, dan memanfaatkan data nyata dari institusi pendidikan.

Decision Tree merupakan salah satu algoritma proses pembuatan keputusan yang memiliki tingkat akurasi yang tinggi dalam mengklasifikasikan data yang besar dibandingkan dengan algoritma pohon keputusan lainnya yang menghasilkan pohon keputusan yang mudah diinterpretasikan [8]. Sistem prediksi dihitung menggunakan metode hitung algoritma *Decision Tree*. *Decision Tree* adalah pemetaan mengenai alternatif pemecahan masalah yang disusun dari masalah yang ditemukan. Metode ini menguraikan faktor-faktor kemungkinan, serta hasil akhir ditunjukkan dengan pohon keputusan [7].

Berdasarkan permasalahan yang ada, maka perlu adanya suatu sistem untuk memprediksi tingkat kelulusan mahasiswa berdasarkan variabel-variabel yang relevan. Selain itu, diperlukan juga untuk algoritma yang cocok sehingga bisa menghasilkan nilai akurasi yang baik. Dengan dibuatnya sistem ini, diharapkan perguruan tinggi bisa membuat kebijakan strategis agar mahasiswa dapat lulus tepat waktu.

Oleh karena itu, penelitian ini bertujuan untuk membangun sistem prediksi kelulusan mahasiswa berbasis *website* dengan memanfaatkan data historis mahasiswa yang diperoleh dari data publik, Kaggle. Dataset ini mencakup berbagai

atribut penting, seperti nama, jenis kelamin, status mahasiswa, umur, status pernikahan, ips1 hingga ips8, ipk, dan status kelulusan. Model klasifikasi akan dikembangkan menggunakan algoritma *Decision Tree* untuk menghasilkan prediksi kelulusan yang akurat. Sistem ini diharapkan dapat membantu institusi dalam pengambilan kebijakan preventif yang efektif guna meningkatkan jumlah lulusan tepat waktu dan memperbaiki akreditasi universitas.

1.2 Pembatasan Masalah

Dalam penelitian ini, penulis memberikan Batasan terhadap ruang lingkup pembahasan guna memastikan penelitian tetap fokus dan sistem dapat dikembangkan secara optimal sesuai dengan tujuan yang ingin dicapai. Adapun batasan-batasan dalam penelitian ini adalah sebagai berikut:

1. Sistem dibangun berbasis *website* menggunakan *framework* Laravel untuk *frontend* dan API Flask sebagai backend untuk menjalankan proses prediksi.
2. Sistem hanya dapat digunakan oleh dua jenis pengguna, yaitu admin dan mahasiswa, dengan fungsi yang berbeda sesuai peran masing-masing.

1.3 Tujuan dan Manfaat

1.3.1 Tujuan

Tujuan penelitian ini adalah pengembangan sebuah sistem atau model yang mampu mengidentifikasi potensi keterlambatan kelulusan mahasiswa secara dini. Deteksi dini ini penting agar pihak akademik dapat melakukan intervensi tepat waktu. Selain itu, sistem yang dikembangkan diharapkan memiliki tingkat kepercayaan (*konfidensi*) yang tinggi, sehingga hasil prediksi dapat diandalkan dan menjadi dasar pengambilan keputusan yang lebih akurat. Dengan demikian, tujuan

ini mendukung efisien proses pendidikan dan membantu peningkatan angka kelulusan tepat waktu di universitas.

1.3.2 Manfaat

Manfaat dikembangkan sistem prediksi kelulusan mahasiswa berbasis *website* dibagi menjadi beberapa bagian antara lain:

1. Mendapatkan peringatan dini untuk pencegahan
Membantu untuk mendeteksi secara dini mahasiswa yang berpotensi terlambat lulus, sehingga dapat dilakukan intervensi lebih cepat.
2. Mendapat bahan pertimbangan untuk menyusun kebijakan preventif
Memberikan data pendukung bagi Program Studi untuk menyusun kebijakan yang dapat mencegah keterlambatan kelulusan.
3. Meminimalkan jumlah mahasiswa yang terlambat lulus
Dengan adanya prediksi dan intervensi dini, jumlah mahasiswa yang tidak lulus tepat waktu dapat ditekan.
4. Mengoptimalkan akreditasi institusi
Tingkat kelulusan tepat waktu yang lebih tinggi berkontribusi langsung terhadap optimalisasi capaian akreditasi institusi.

1.4 Tinjauan Pustaka

Sebuah penelitian menerapkan algoritma *Decision Tree* untuk klasifikasi prediksi kelulusan mahasiswa menggunakan data 1.739 alumni dari beberapa perguruan tinggi di Palembang. Hasilnya menunjukkan akurasi 87,92%, *precision* 94,87%, dan *recall* 81,32%. Penelitian ini unggul dalam penggunaan data besar dan

validasi *K-fold*, namun belum membandingkan dengan algoritma lain maupun pengembangan sistem prediksi terintegrasi [9].

Penelitian lain menyatakan bahwa algoritma *Decision Tree* dapat digunakan untuk memprediksi tingkat kelulusan mahasiswa PTIK UIN Bukittinggi berdasarkan data akademik, demografis, dan sosial ekonomi, dengan akurasi mencapai 85,37%. Penelitian ini telah menerapkan *cross-validation* untuk meningkatkan keandalan model, namun belum mengembangkan sistem prediksi terintegrasi [10].

Sebuah penelitian menunjukkan bahwa algoritma *Decision Tree* mampu memprediksi kelulusan mahasiswa STT Ronggolawe Cepu dengan akurasi 91%, *precision* 0,92%, dan *recall* 0,91%. Penelitian ini menggunakan atribut akademik seperti umur, jenis kelamin, jurusan, SKS, IP, dan beasiswa serta menerapkan SMOTE untuk mengatasi ketidakseimbangan data, namun belum membandingkan dengan algoritma lain maupun mengembangkan sistem prediksi terintegrasi [11].

Selain itu, penelitian lain juga menunjukkan bahwa algoritma C4.5 dapat memprediksi kelulusan mahasiswa STIA Trintis Ambon menggunakan atribut SKS dan IPK, dengan akurasi 85%, *precision* 92%, dan *recall* 93,2%. Model ini mampu mengidentifikasi faktor utama yang mempengaruhi kelulusan tepat waktu, namun belum membandingkan dengan algoritma lain maupun mengembangkan sistem prediksi terintegrasi [12].

Penelitian lain membandingkan algoritma *Decision Tree* C4.5 dan *Naïve Bayes* untuk memprediksi kelulusan mahasiswa tepat waktu menggunakan 10 atribut akademik seperti IP semester dan umur. Hasilnya menunjukkan bahwa

Decision Tree memiliki akurasi lebih tinggi (88,92%) dibandingkan *Naïve Bayes* (84,98%), sehingga dinilai lebih efektif dalam meningkatkan ketepatan prediksi kelulusan [13].

Selain itu, penelitian lain menunjukkan bahwa algoritma *Decision Tree* C4.5 dapat menganalisis kelulusan mahasiswa D3 Manajemen Informatika Universitas Muhammadiyah Pekajangan Pekalongan dengan akurasi 73,48%. Dengan atribut tahun masuk menjadi faktor dominan yang mempengaruhi kelulusan, sedangkan penelitian ini masih terbatas pada satu algoritma dan belum melibatkan faktor non-akademik [14].

Untuk melihat tabel penelitian terdahulu dapat dilihat pada Tabel 1.1.

Tabel 1.1. Penelitian Terdahulu

No	Judul	Hasil	Pembeda
1	Qisthiano, <i>et al.</i> – Penerapan <i>Algoritma</i> <i>Decision Tree</i> dalam Klasifikasi Data Prediksi Kelulusan Mahasiswa	Algoritma <i>Decision Tree</i> mencapai akurasi 87,93%, <i>precision</i> 94,87%, dan <i>recall</i> 81,32% dalam memprediksi kelulusan 1.739 mahasiswa di Palembang, namun belum terintegrasi dalam sistem prediksi.	Fokus pada klasifikasi kelulusan, sedangkan penelitian Predict fokus pada klasifikasi dan pengembangan sistem prediksi terintegrasi.

No	Judul	Hasil	Pembeda
2	Yuspita, <i>et al.</i> – Penerapan Algoritma Klasifikasi Tingkat Kelulusan Mahasiswa Menggunakan Rapidminer	Algoritma <i>Decision Tree</i> menghasilkan akurasi 85,37% dalam memprediksi kelulusan mahasiswa berdasarkan data akademik, demografis, dan sosial ekonomi dari 164 mahasiswa.	Menggunakan data <i>e-campus</i> nyata dan <i>cross-validation</i> , sedangkan sistem Predict menggunakan data Kaggle dengan hasil akurasi sebesar 88,16% serta dilengkapi sistem prediksi yang terintegrasi berbasis <i>web</i> yang mendukung pengambilan keputusan perguruan tinggi.
3	Berlian, <i>et al.</i> – <i>Penerapan Machine Learning</i> untuk Memprediksi Kelulusan Mahasiswa Menggunakan Algoritma <i>Decision Tree</i>	Akurasi 91%, <i>precision</i> 0,92%, <i>recall</i> 0,91%; model mampu menangani ketidakseimbangan data dengan SMOTE dan menghasilkan akurasi tinggi.	Menggunakan data sampel terbatas dan belum berbasis sistem terintegrasi, sedangkan penelitian Predict mengembangkan sistem prediksi berbasis <i>web</i> dengan data Kaggle yang lebih luas, dilengkapi fitur <i>report</i> data pada admin serta filter berdasarkan tahun dan program studi.

No	Judul	Hasil	Pembeda
4	Daniel Hartono Renyut, <i>et al.</i> – Prediksi Kelulusan Mahasiswa Menggunakan Algoritma C.45 (Studi Kasus, Sekolah Tinggi Ilmu Administrasi Trintis Ambon)	Algoritma C4.5 menghasilkan akurasi 85%, <i>precision</i> 92%, dan <i>recall</i> 93,2%, menunjukkan model cukup baik dalam klasifikasi kelulusan tepat waktu.	Menggunakan data asli institusi dengan hasil akurat melalui pohon keputusan, namun belum memiliki implementasi sistem terintegrasi berbasis <i>web</i> seperti pada penelitian Predict yang dilengkapi fitur prediksi keseluruhan mahasiswa pada admin.
5	Wahyudi, <i>et al.</i> – Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan <i>Metode</i> <i>Decision Tree</i> dan <i>Naïve</i> <i>Bayes</i>	Algoritma <i>Decision</i> <i>Tree</i> C4.5 menghasilkan akurasi 88,92%, lebih tinggi dibandingkan <i>Naïve</i> <i>Bayes</i> sebesar 84,98%, menggunakan 10 atribut akademik dan <i>validasi K-Fold</i> .	Hanya berupa eksperimen algoritma tanpa implementasi sistem, sedangkan penelitian Predict mengembangkan sistem prediksi yang dilengkapi fitur prediksi bagi mahasiswa dan dapat melihat hasil riwayat prediksinya.

No	Judul	Hasil	Pembeda
6	Fatkhudin, <i>et al.</i> – Algoritma <i>Decision Tree</i> C.45 Dalam Analisis Kelulusan Mahasiswa Program Studi Manajemen Informatika UMPP	Algoritma <i>Decision Tree</i> C4.5 menghasilkan akurasi 73,48%, dengan tahun masuk sebagai faktor utama yang mempengaruhi kelulusan mahasiswa.	Hanya menggunakan variabel akademik terbatas, sedangkan penelitian Predict menambah jumlah variabel akademik dan non-akademik.

Berdasarkan tinjauan pustaka terhadap penelitian terdahulu, penelitian ini bertujuan untuk mengembangkan sistem prediksi kelulusan mahasiswa berbasis *website* yang terintegrasi dan mudah digunakan. Penelitian sebelumnya menunjukkan bahwa algoritma *Decision Tree* mampu memberikan hasil prediksi dengan tingkat akurasi yang tinggi, namun sebagian besar masih terbatas pada tahap eksperimen dan belum diterapkan dalam bentuk sistem yang dimanfaatkan langsung oleh perguruan tinggi. Oleh karena itu, penelitian ini melakukan pengembangan lebih lanjut dengan membangun sistem berbasis *website* yang dilengkapi dengan fitur prediksi bagi mahasiswa dan admin. Mahasiswa dapat melakukan prediksi secara mandiri serta melihat riwayat hasil prediksinya,

sedangkan admin dapat melakukan prediksi secara keseluruhan, memantau hasil prediksi per tahun dan per program studi, serta menghasilkan laporan (*report*) data. Dengan adanya sistem ini, perguruan tinggi diharapkan dapat lebih mudah dalam memantau dan mengevaluasi potensi kelulusan mahasiswa setiap tahunnya.

1.5 Data Penelitian

Penelitian ini menggunakan dataset data historis mahasiswa yang diperoleh dari Kaggle (<https://www.kaggle.com/datasets/hafizhathallah/kelulusan-mahasiswa/data>). Dataset ini memuat informasi lengkap yang relevan untuk menganalisis serta memprediksi kelulusan mahasiswa. Data yang digunakan mencakup atribut-atribut penting yang berhubungan dengan informasi demografi maupun riwayat akademik mahasiswa. Atribut-atribut tersebut kemudian dijadikan dasar dalam membangun model prediksi untuk memprediksi kelulusan mahasiswa.

1.5.1 Data Historis Mahasiswa

Dataset mahasiswa yang digunakan dalam penelitian ini memiliki atribut yang menggambarkan karakteristik serta performa akademik mahasiswa. Atribut Nama berisi nama lengkap mahasiswa, Jenis Kelamin menunjukkan kategori jenis kelamin mahasiswa yaitu Laki-laki atau Perempuan. Atribut Status Mahasiswa memuat informasi mengenai status pekerjaan mahasiswa selama menempuh pendidikan, misalnya bekerja atau hanya berstatus mahasiswa.

Selanjutnya, terdapat atribut umur yang berisi usia mahasiswa, serta status nikah yang menunjukkan status pernikahan mahasiswa dengan kategori belum menikah atau menikah. Untuk mengukur performa akademik, dataset ini juga memuat Indeks Prestasi Semester (IPS) mulai dari semester 1 hingga semester 8.

Dari seluruh nilai IPS tersebut kemudian dihitung Indeks Prestasi Kumulatif (IPK) sebagai nilai rata-rata keseluruhan akademik mahasiswa.

Atribut akhir adalah status kelulusan yang menjadi variabel target dalam penelitian ini, dengan kategori tepat waktu atau terlambat. Atribut ini digunakan sebagai acuan dalam proses analisis dan prediksi kelulusan mahasiswa.

1.5.2 Alat Penelitian

Penelitian ini menggunakan berbagai perangkat keras dan perangkat lunak untuk mendukung proses pengumpulan data, pemodelan, dan pengembangan aplikasi. Berikut adalah tabel yang mencakup alat penelitian yang digunakan disajikan pada Tabel 1.2.

Tabel 1.2. Alat Penelitian

No	Perangkat Keras Alat Penelitian	
	Nama Alat	Fungsi
1.	Laptop dengan spesifikasi mendukung	Sebagai perangkat utama untuk mengembangkan, menguji, dan menjalankan aplikasi.
2.	Ruang Penyimpanan (SSD)	Sebagai komponen penyimpanan untuk menyimpan data, aset proyek, source code, serta file model dan dataset.
3.	RAM minimum 4GB	Mendukung proses komputasi dan pemrosesan data secara efisien selama pengembangan aplikasi.

No	Perangkat Lunak Alat Penelitian	
	Nama Alat	Fungsi
1.	Windows OS	Sebagai operasi utama pada laptop/komputer pengembang.
2.	Google Colaboraty	Lingkungan pemrograman berbasis cloud untuk pemrosesan dataset dan pelatihan model <i>decision Tree</i> menggunakan <i>python</i> .
3.	Visual Studio Code	Sebagai aplikasi <i>code</i> editor yang digunakan untuk membuat aplikasi.
4.	<i>MySQL</i>	Digunakan sebagai sistem manajemen basis data untuk menyimpan data mahasiswa dan hasil prediksi.
5.	Mozilla Firefox/Chrome	Digunakan untuk menguji dan menampilkan antarmuka aplikasi berbasis <i>web</i> .
6.	Flask	Digunakan untuk menghubungkan model prediksi dengan antarmuka <i>web</i> , serta menangani proses <i>routing</i> , <i>input</i> data, dan penampilan hasil prediksi kepada pengguna.
7.	Pandas	<i>Library</i> Python untuk manipulasi dan analisis data.

8.	Python	Python digunakan untuk proses pengolahan data, pelatihan model <i>machine learning</i> (<i>Decision Tree</i>).
----	--------	--