

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pendidikan tinggi di Indonesia menghadapi tantangan serius dalam memastikan mahasiswa menyelesaikan studi tepat waktu. Tingkat kelulusan mahasiswa dalam periode yang telah ditentukan bukan hanya menjadi tolok ukur keberhasilan individu, tetapi juga mencerminkan kualitas institusi pendidikan itu sendiri. Menurut Badan Akreditasi Nasional Perguruan Tinggi (BAN-PT), persentase mahasiswa yang lulus tepat waktu merupakan salah satu indikator utama dalam Instrumen Pemantauan dan Evaluasi Peringkat Akreditasi (IPEPA). Dalam Peraturan BAN-PT Nomor 23 Tahun 2022, kelulusan tepat waktu menjadi bagian krusial dalam menilai kualitas suatu perguruan tinggi (BAN-PT, 2022). Perguruan tinggi yang gagal menjaga rasio kelulusan tepat waktu berisiko mengalami dampak negatif, mulai dari penurunan peringkat akreditasi hingga menurunnya daya saing di tingkat nasional maupun internasional. Oleh karena itu, diperlukan solusi yang inovatif dan berbasis teknologi untuk mengatasi permasalahan ini.

Sebagai salah satu politeknik swasta terkemuka di Riau, Politeknik Caltex Riau (PCR) menghadapi tantangan nyata dalam memastikan kelulusan mahasiswa tepat waktu. Berdasarkan data akademik PCR dalam lima tahun terakhir, 12,41% mahasiswa gagal menyelesaikan studi dalam waktu yang ditetapkan, yaitu empat tahun. Dalam program D4, mahasiswa diwajibkan menyelesaikan 148 SKS yang telah dipaketkan dalam kurikulum. Namun, faktor-faktor seperti prestasi akademik, jalur penerimaan mahasiswa baru (PMB), serta kondisi sosial ekonomi sering kali menjadi penghambat utama dalam penyelesaian studi. Tanpa adanya sistem pemantauan yang efektif, sulit bagi pihak akademik untuk mengidentifikasi mahasiswa yang berisiko mengalami keterlambatan dan melakukan intervensi yang tepat waktu.

Penelitian sebelumnya telah menunjukkan bahwa Indeks Prestasi Kumulatif (IPK) memiliki korelasi kuat dengan kelulusan tepat waktu (Angelia et al., 2023). Selain itu, faktor sosial ekonomi, seperti pendapatan orang tua, juga memainkan peran penting dalam pencapaian akademik mahasiswa (Jaya, 2016). Sayangnya, pendekatan konvensional dalam mengidentifikasi mahasiswa yang berisiko terlambat lulus sering kali bersifat reaktif, yakni hanya dilakukan setelah

mahasiswa mengalami kesulitan yang signifikan. Oleh karena itu, diperlukan pendekatan yang lebih prediktif dan berbasis data untuk mengatasi permasalahan ini sebelum dampaknya semakin besar.

Seiring dengan kemajuan teknologi, *Machine Learning* telah terbukti menjadi metode yang efektif dalam memprediksi keberhasilan akademik mahasiswa. Algoritma seperti *Logistic Regression*, *Support Vector Machines* (SVM), *K-Nearest Neighbors* (KNN), *Decision Tree*, dan *Random Forest* telah banyak digunakan dalam penelitian untuk mengolah data akademik, demografi, serta faktor sosial ekonomi guna menghasilkan prediksi yang lebih akurat dibandingkan dengan metode konvensional (Riska Rismaya et al., 2025). Dengan teknologi ini, institusi pendidikan dapat mengantisipasi keterlambatan studi mahasiswa sejak dini dan mengambil langkah intervensi yang lebih cepat serta tepat sasaran.

Oleh karena itu, pengembangan sistem berbasis *website* yang memanfaatkan *machine learning* menjadi kebutuhan bagi PCR. Sistem ini memberikan visualisasi prediksi keterlambatan studi secara interaktif, sehingga pihak akademik dapat dengan cepat mengidentifikasi mahasiswa yang berisiko dan mengambil tindakan yang diperlukan. Dengan solusi ini, PCR berhasil meningkatkan efektivitas manajemen akademik, memastikan lebih banyak mahasiswa lulus tepat waktu, meningkatkan reputasi institusi, serta memperkuat daya saing di tingkat nasional maupun global.

1.2 Perumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, permasalahan utama yang diangkat dalam penelitian ini adalah:

- 1) Algoritma *Machine Learning* apa yang paling optimal dalam memprediksi keterlambatan mahasiswa Politeknik Caltex Riau (PCR) dalam menyelesaikan studi?
- 2) Bagaimana cara mengembangkan sistem berbasis *web* yang dapat menampilkan hasil prediksi keterlambatan mahasiswa dalam menyelesaikan studi?

1.3 Batasan Masalah

Penelitian ini dibatasi pada

- 1) Penelitian ini hanya bertujuan untuk memprediksi keterlambatan mahasiswa Politeknik Caltex Riau (PCR) dalam menyelesaikan studi menggunakan algoritma *Machine Learning* berbasis *supervised learning*.
- 2) Model yang dibangun menggunakan 5 (lima) algoritma *machine learning*, yaitu *Logistic Regression*, *Support Vector Machines* (SVM), *K-Nearest Neighbors* (KNN), *Decision Tree*, dan *Random Forest*. Hanya model dengan performa terbaik yang akan diimplementasikan dalam sistem berbasis *website* untuk memastikan prediksi yang lebih akurat.
- 3) Data yang digunakan dalam penelitian ini mencakup informasi akademik dan demografis mahasiswa PCR, seperti program studi, tahun masuk, jenis kelamin, IPK, IPS, serta status penerimaan dan data akademik lainnya.
- 4) *Dataset* yang digunakan bersifat privat dan tidak akan dibagikan kepada publik untuk menjaga kerahasiaan informasi mahasiswa.
- 5) Sistem yang dikembangkan hanya berbasis *web* untuk menampilkan hasil prediksi kelulusan tepat waktu, tanpa diintegrasikan dengan sistem manajemen akademik lainnya.
- 6) Penelitian ini berfokus pada pengembangan model prediksi dan tidak mencakup evaluasi terhadap faktor-faktor eksternal, seperti kebijakan institusi atau perubahan kurikulum, yang mungkin memengaruhi masa studi mahasiswa.
- 7) Model prediksi ini hanya akan digunakan secara internal oleh pihak kampus dan tidak bersifat publik.

1.4 Tujuan dan Manfaat

1.4.1 Tujuan

Penelitian ini bertujuan untuk menerapkan *Machine Learning* dalam memprediksi keterlambatan mahasiswa Politeknik Caltex Riau (PCR) dalam menyelesaikan studi serta mengembangkan sistem berbasis *web* untuk memvisualisasikan hasil prediksi tersebut, sekaligus mengevaluasi performa berbagai algoritma *Machine Learning* yang digunakan.

1.4.2 Manfaat

Manfaat yang diharapkan dari penelitian ini adalah:

- 1) Membantu pihak kampus dalam memonitor mahasiswa yang berisiko terlambat lulus dengan menggunakan teknologi prediksi berbasis *machine learning*.
- 2) Menyediakan alat berbasis *web* yang dapat digunakan oleh manajemen akademik PCR untuk pengambilan keputusan yang lebih efektif.
- 3) Memberikan wawasan tentang pemanfaatan teknologi *Machine Learning* dalam pengelolaan akademik untuk meningkatkan efisiensi dan kualitas pendidikan.

1.5 Metodologi Penelitian

Metode penelitian yang digunakan dalam penelitian ini meliputi:

1) Identifikasi Masalah

Penelitian dimulai dengan mengidentifikasi alasan utama mengapa banyak mahasiswa tidak lulus tepat waktu. Langkah ini mencakup analisis kondisi akademik mahasiswa di Politeknik Caltex Riau dan eksplorasi potensi *Machine Learning* sebagai solusi untuk mengatasi permasalahan ini.

2) Studi Literatur

Studi literatur dilakukan dengan meninjau penelitian terdahulu yang relevan, jurnal ilmiah, dan publikasi lainnya terkait penerapan *Machine Learning* dalam

pendidikan tinggi. Literatur ini menjadi dasar untuk memahami metode terbaik yang dapat digunakan.

3) Pengumpulan Data

Data yang digunakan dalam penelitian ini diambil dari *dataset* akademik mahasiswa PCR selama lima tahun terakhir, mencakup informasi seperti program studi, tahun masuk, jenis kelamin, IPK, IPS, status penerimaan KPS, dan faktor akademik lainnya. Langkah ini melibatkan proses pembersihan data, seperti menangani nilai yang hilang, normalisasi data, dan memastikan data yang digunakan valid untuk analisis.

4) Perancangan

Perancangan model prediksi dilakukan menggunakan lima algoritma *machine learning*, yaitu *Logistic Regression*, *Support Vector Machines (SVM)*, *K-Nearest Neighbors (KNN)*, *Decision Tree*, dan *Random Forest*. Model-model ini akan diuji untuk menentukan algoritma terbaik dalam memprediksi kelulusan tepat waktu mahasiswa, dengan mempertimbangkan faktor-faktor akademik dan demografis yang tersedia.

5) Implementasi

Algoritma dengan performa terbaik akan diterapkan sebagai metode prediksi, dilatih menggunakan *dataset* yang telah dikumpulkan untuk menghasilkan model prediksi. Selain itu, sistem berbasis *website* dikembangkan sesuai desain, mencakup fitur pengumpulan data, pengolahan, dan visualisasi hasil prediksi.

6) Pengujian

Pengujian model dilakukan menggunakan metrik evaluasi seperti akurasi, *precision*, *recall*, dan *F1-score*. Pengujian ini bertujuan untuk menilai performa model dalam memprediksi keterlambatan mahasiswa Politeknik Caltex Riau (PCR) dalam menyelesaikan studi.

7) Analisis dan Penarikan Kesimpulan

Hasil pengujian model dievaluasi untuk membandingkan performa algoritma yang digunakan. Evaluasi ini juga

mencakup analisis potensi pengembangan model dan penerapannya dalam sistem akademik PCR.

1.6 Sistematika Penulisan

Sistematika penulisan laporan proyek akhir ini secara keseluruhan terdiri dari empat bab, masing-masing terdiri dari beberapa sub bab. Adapun pokok pembahasan dari masing-masing bab tersebut secara garis besar sebagai berikut:

BAB I PENDAHULUAN

Bab ini menguraikan tentang latar belakang masalah, perumusan masalah dan ruang lingkup masalah, tujuan dan manfaat penelitian, metodologi penelitian dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Bab ini menguraikan beberapa hasil penelitian terdahulu dan landasan teori yang diperlukan untuk merancang sistem.

BAB III PERANCANGAN

Bab ini menjelaskan tentang perancangan sistem terdiri dari perancangan sistem yang akan dibangun.

BAB IV PENGUJIAN DAN ANALISIS

Bab ini berisi informasi mengenai hasil implementasi dan pengujian serta analisis pada aplikasi yang telah dibangun sesuai dengan tahapan perancangan.

BAB V KESIMPULAN DAN SARAN

Bab ini mnenguraikan kesimpulan dari pengerjaan proyek akhir dan saran yang bisa diterapkan kedepannya.