

BAB 1

PENDAHULUAN

1.1. Latar Belakang

Pendidikan tinggi seperti yang tertuang dalam undang-undang Republik Indonesia nomor 12 tahun 2012 tentang pendidikan tinggi mempunyai peran yang strategis dalam mencerdaskan kehidupan bangsa untuk meningkatkan daya saing bangsa dalam menghadapi globalisasi di segala bidang. Tentunya sebuah perguruan tinggi harus terus berupaya menjaga mutu pendidikan dengan meningkatkan kualitas pembelajaran dan layanan akademik untuk mahasiswa (Salmu and Solichin, 2017). Aspek penting yang harus diperhatikan dalam evaluasi program studi dapat dilihat dari lama studi mahasiswa sebagai salah satu indikator keberhasilan sebuah program studi (Etriyanti, Syamsuar and Kunang, 2020). Mengingat kelulusan merupakan salah satu bidang yang termasuk ke dalam standar penjamin mutu internal suatu perguruan tinggi (Widaningsih, 2019).

Banyaknya jumlah mahasiswa tidak menjamin bagusnya akreditasi suatu prodi (Saputro and Sari, 2019). Hal ini disebabkan adanya salah satu instrumen dalam akreditasi menurut Badan Akreditasi Nasional Perguruan Tinggi (BAN-PT) pada Peraturan Badan Akreditasi Nasional perguruan Tinggi Nomor 2 Tahun 2017 tentang Sistem Akreditasi Nasional Pendidikan Tinggi yang mempengaruhi akreditasi adalah dampak, capaian, mutu serta produktivitas luaran yang bermanfaat bagi masyarakat sehingga Program studi mempunyai kewajiban untuk memantau perkembangan studi dari mahasiswanya dan memprediksi lama studi

dari setiap mahasiswanya guna menentukan dan mengantisipasi terjadinya penumpukan jumlah mahasiswa yang akan menyebabkan kinerja program studi menjadi kurang baik (Jananto, 2013).

Sekolah Tinggi Agama Islam Negeri (STAIN) merupakan perguruan tinggi negeri satu-satunya di Kota Metro yang berdiri berdasarkan SK Presiden No.11 tahun 1997 dengan ditetapkan nama STAIN Jurai Siwo Metro dan pada Tahun 2016 adalah tahun peralihan STAIN menjadi IAIN yang tertuang dalam Peraturan Presiden No.71 tanggal 1 Agustus 2016. IAIN Metro memiliki 6 fakultas dengan jumlah mahasiswa 8.271 yang terbagi menjadi 4 fakultas untuk program sarjana sebanyak 7.940 mahasiswa dan 2 program pasca sarjana sebanyak 391 mahasiswa. Sumber data dari sistem informasi akademik IAIN Metro tahun akademik 2021/2022 menyebutkan fakultas ekonomi dan bisnis islam memiliki jumlah mahasiswa terbanyak dari jumlah mahasiswa pada fakultas yang lain yaitu 2.628 mahasiswa dengan jumlah mahasiswa terbanyak pada prodi ekonomi syariah dan perbankan syariah dengan total masing-masing mahasiswa 1.228 mahasiswa dan 871 mahasiswa.

Merujuk pada buku pedoman akademik IAIN Metro tahun akademik 2020/2021 menyebutkan untuk program S1 memiliki beban studi sekurang-kurangnya 144 sks dan sebanyak-banyaknya 160 sks dengan ditempuh selama 8 semester atau paling lama 7 tahun atau 14 semester. Sementara Data kelulusan yang diperoleh dari bagian akademik Fakultas Ekonomi dan bisnis untuk program sarjana diketahui lama studi dari mahasiswa sebagian besar berada di atas 4

Tahun. Data lama studi mahasiswa Fakultas Ekonomi dan Bisnis Islam IAIN Metro dapat dilihat pada tabel pada Tabel 1.1 di bawah ini:

Tabel 1.1 Lama Studi Alumni Mahasiswa Fakultas Ekonomi dan Bisnis Islam IAIN Metro

Lama studi	Jumlah kelulusan	Persentase kelulusan
7 tahun	43	10%
6 tahun	118	27,44%
5 tahun	223	51,86%
4 tahun	46	10,70%
Σ	430	100%

Sumber : Data yang diolah 2021

Berdasarkan persentase jumlah lulusan tahun akademik 2020/2021 dapat dilihat bahwa hanya ada 10,70% mahasiswa yang dapat mencapai target lulus tepat waktu pada semester 8 atau 4 tahun masa studi. Melihat permasalahan tersebut maka program studi menghadapi permasalahan yang muncul akibat adanya keterlambatan target lulus mahasiswa yang sudah ditentukan oleh institusi sehingga akan berdampak pada penumpukan jumlah mahasiswa. Upaya dalam melakukan prediksi dan analisis kinerja mahasiswa merupakan tonggak penting dalam lingkungan pendidikan untuk membangun masa depan peserta didik (Kaur, Singh and Singh, 2015). Penelitian ini penting dilakukan sebagai langkah awal bagi prodi untuk mengetahui masa studi mahasiswa lebih dini maka pihak jurusan dapat melakukan tindakan-tindakan strategis supaya mahasiswa dapat lulus tepat waktu sekaligus meningkatkan kualitas jurusan itu sendiri (Ginting, Zarman and Hamidah, 2014).

Salah satu teknik untuk melakukan prediksi adalah dengan menggunakan penggalian data atau data mining (Salmu and Solichin, 2017). Fungsi data mining dapat digunakan untuk memvalidasi temuan dengan menerapkan pola yang terdeteksi ke subset data baru dan digunakan di berbagai bidang misalnya perbankan, keuangan, penjualan eceran, kesehatan dan pendidikan untuk mendukung organisasi dalam menemukan cara untuk meningkatkan kinerja mereka (Ramaswami *et al.*, 2019). Ada berbagai teknik dan alat penambangan data yang telah dikembangkan untuk mengubah volume data yang terus bertambah menjadi informasi yang berharga dan memberikan potensi yang signifikan bagi organisasi maupun peneliti untuk menemukan informasi implisit di berbagai bidang termasuk pendidikan serta mampu menjelaskan penyebab masalah lama masa studi yang dihadapi oleh mahasiswa, termasuk tingkat kegagalan belajar atau drop out dan kinerja akademik yang buruk (Badr *et al.*, 2016).

Aplikasi untuk menggali data dalam konteks pendidikan telah mengalami pertumbuhan yang cepat (Altujjar *et al.*, 2016). Sejumlah penelitian baru-baru ini telah menerapkan teknik data mining pada bidang pendidikan untuk mengklasifikasikan dan memprediksi kinerja siswa di berbagai lembaga pendidikan (Mohamed, Rizaner and Hakan, 2016). Penggunaan teknik data mining yang digunakan untuk mengekstrak informasi penting dari kumpulan data yang dihasilkan dalam konteks pendidikan melalui proses pembelajaran mesin yang telah diterapkan untuk mempelajari dan memprediksi kinerja siswa disebut *Education Data Mining* (Israel and Salau, 2019). *Education data mining*

merupakan bidang baru yang muncul dan diterapkan pada sektor pendidikan (Kaur, Singh and Singh, 2015). *Education data mining* mengacu pada teknik data mining yang digunakan untuk menganalisis data pendidikan (Pallathadka *et al.*, 2021).

Education Data Mining memberikan kesempatan kepada institusi pendidikan untuk memanfaatkan wawasan berbasis data dan membangun proses operasional yang lebih efisien (Ramaswami *et al.*, 2019). Informasi yang dihasilkan *Education Data Mining* dapat bermanfaat bagi beberapa pemangku kepentingan di bidang pendidikan seperti membantu akademik mengevaluasi struktur pembelajaran dan strategi pengajaran, memberikan rekomendasi pembelajaran berdasarkan kemajuan mahasiswa serta untuk memprediksi kinerja mahasiswa (Altujjar *et al.*, 2016). *Education data mining* (EDM) memberikan kesempatan kepada institusi pendidikan untuk memanfaatkan wawasan berbasis data dan membangun proses operasional yang lebih efisien (Ramaswami *et al.*, 2019).

EDM menggunakan pembelajaran mesin dan teknik penambangan data ini untuk mengeksplorasi data dari pengaturan pendidikan untuk mengetahui prediksi dan pola yang menjadi ciri perilaku dan kinerja siswa (Hai, Zhang and Zhang, 2017). Melalui informasi dan pola yang berguna dapat digunakan dalam memprediksi kinerja siswa sehingga dapat membantu para pendidik dalam memberikan pendekatan pengajaran yang efektif (Shahiri, Husain and Nur'aini Abdul Rashid, 2015). Salah satu pendekatan yang diterapkan dalam *Educational Data mining* (EDM) adalah model prediktif (Widyahastuti and Tjhin, 2018). ada

beberapa tugas yang digunakan untuk membangun pemodelan prediktif yaitu klasifikasi, regresi dan kategorisasi dan yang paling populer untuk memprediksi kinerja siswa adalah klasifikasi (Shahiri, Husain and Nur'aini Abdul Rashid, 2015). Data mining pendidikan menggunakan banyak teknik seperti Decision Trees, Neural Networks, Naïve Bayes, K-nearest neighbour dan banyak lainnya (Kaur, Singh and Singh, 2015).

Ada beberapa penelitian terdahulu yang menggunakan teknik data mining pada bidang pendidikan untuk membantu memprediksi kelulusan mahasiswa antara lain Etriyanti, Samsuar dan Kunang (2020) memprediksi kelulusan mahasiswa menggunakan dua algoritma yaitu algoritma C4.5 dan Naïve Bayes diperoleh hasil akhir bahwa algoritma C4.5 memiliki tingkat akurasi 79,08% lebih tinggi dari algoritma naïve bayes yaitu 78,46% (Etriyanti, Syamsuar and Kunang, 2020). Hasil penelitian Budiantara., et al (2020) membandingkan tiga algoritma yaitu Decision tree, Naïve Bayes dan K-Nearest Neighbor (K-NN) untuk memprediksi mahasiswa lulus tepat waktu dengan hasil akurasi yang diperoleh dari hasil Decision tree (C4.5) sebesar 98,04%, algoritma naïve bayes 96,00% dan K-NN sebesar 90,00% (Budiyantara *et al.*, 2020). Fahri (2020) menggunakan algoritma C4.5, algoritma ID3 dan KNN untuk memprediksi tingkat kelulusan mahasiswa dengan hasil akurasi yang diperoleh untuk model algoritma C4.5 sebesar 92,20% (Fahri, 2018).

Sementara penelitian Verma, Yadav dan Kholiya (2022) memprediksi Prestasi Akademik Mahasiswa Teknik dengan Menggunakan Teknik Data Mining Decision Tree, k-NN, SVM, Random Forest dan Naïve Bayes, diperoleh nilai

akurasi tertinggi adalah Naïve Bayes dengan akurasi sekitar 89% (Verma, Yadav and Kholiya, 2022). Hasil penelitian yang sama juga diperoleh dari penelitian Al-Shabandar., et al (2017) mengimpengimplementasikan berbagai model *Machine Learning linear* dan *nonlinear* telah antara lain *Logistic Regression (LR)*, *Linear Discriminant Analysis (LDA)*, *Naive Bayes (NB)*, *Support Vector Machine (SVM)*, *Decision Tree (DT)*, *Random Forest (RF)*, *Neural Network (MLP)* dan *Self Organized Map (SOM)* dengan hasil menunjukkan bahwa Random Forest kinerja tertinggi dari model yang diuji.yaitu diperoleh nilai *Area Under Curve (AUC)* sebesar 0,90 (Al-Shabandar *et al.*, 2017). Sedangkan Hasil penelitian Ramaswarni.,at al (2019) yang memprediksi prestasi akademik siswa dengan menggunakan empat algoritma yaitu Naïve Bayes, Random Forest, Statistik Logistik dan k-NN dengan hasil menunjukkan bahwa Random Forest lebih akurat daripada tiga algoritma lainnya dengan cara yang signifikan secara statistik yaitu 88% (Ramaswami *et al.*, 2019).

Berdasarkan uraian latar belakang dan penelitian sebelumnya maka Teknik data mining yang akan digunakan dalam penelitian ini adalah algoritma Algoritma C4.5 dan Algoritma Random Forest untuk membandingkan akurasi teknik EDM yang ada dalam mengidentifikasi awal bagi mahasiswa yang cenderung gagal dengan resiko tinggi dengan mendapatkan tingkat akurasi yang lebih tinggi.

1.2. Ruang lingkup masalah

Ruang lingkup masalah dalam penelitian ini dibatasi pada masalah perbandingan algoritma Algoritma C4.5 dan Algoritma Random Forest pada masa studi mahasiswa dengan ketentuan sebagai berikut:

- 1) Data yang digunakan adalah data mahasiswa Fakultas Ekonomi dan Bisnis Islam IAIN Metro
- 2) Dataset menggunakan data akademik mahasiswa angkatan 2017 sampai 2018 yang dinyatakan lulus Tahun 2021-2022.
- 3) *Tools* yang digunakan dalam penelitian ini menggunakan Rapid Miner9

1.3. Rumusan masalah

Rumusan masalah penelitian ini adalah bagaimana perbandingan tingkat performa algoritma C4.5 dan algoritma Random Forest dalam memprediksi masa studi mahasiswa.

1.4. Tujuan Penelitian

Tujuan dalam penelitian ini adalah untuk membandingkan kinerja algoritma yang memiliki tingkat akurasi tinggi dengan menggunakan algoritma C4.5 dan algoritma Random Forest pada masa studi mahasiswa

1.5. Manfaat penelitian

Memberikan kontribusi dalam penerapan metode data mining untuk memprediksi kelulusan mahasiswa dan institusi dalam menentukan strategi

dengan memberikan perhatian lebih bagi mahasiswa yang diprediksi akan terlambat lulus sebagai bentuk upaya untuk mewujudkan kualitas sebuah perguruan tinggi pada aspek lulusan.

1.6. Sistematika Penulisan

Sistematika dalam penulisan penelitian ini adalah sebagai berikut :

BAB I PENDAHULUAN

Bab ini menguraikan secara umum meliputi latar belakang, perumusan masalah, ruang lingkup penelitian, tujuan dan manfaat dan sistematika penulisan.

BAB II LANDASAN TEORI

Bab ini memuat tentang teori-teori yang mendukung penelitian yang dilakukan oleh peneliti meliputi konsep data mining dan teknologi yang menunjang untuk menguji perbandingan algoritma Naive Bayes dan algoritma C4.5 dalam penelitian ini.

BAB III METODOLOGI PENELITIAN

Bab ini berisi metode-metode penelitian yang digunakan serta kebutuhan sistem aplikasi yang digunakan untuk membandingkan dua algoritma dalam memprediksi kelulusan mahasiswa baik berupa analisis kebutuhan data, desain dan rancangan konteks diagram, rancangan *database*, dan rancangan tampilan.

BAB IV HASIL PENELITIAN DAN PEMBAHASAN

Bab ini berisi pengujian Algoritma C4.5 dan AlgoritmaRandom Forest pada masa studi mahasiswa dan juga berisi tampilan aplikasi beserta penjelasannya.

BAB V SIMPULAN DAN SARAN

Bab ini berisi tentang kesimpulan dari hasil penelitian dan juga saran-saran yang ditujukan kepada semua pihak yang bersangkutan.

DAFTAR PUSTAKA

LAMPIRAN