

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perguruan tinggi merupakan suatu pihak yang menyelenggarakan atau melakukan pendidikan akademik bagi para mahasiswa [1]. Berdasarkan pengelolanya, perguruan tinggi terbagi menjadi 3 jenis yaitu Perguruan Tinggi Negeri (PTN), Perguruan Tinggi Swasta (PTS), dan Perguruan Tinggi Kedinasan (PTK). Perbedaan jenis perguruan tinggi ini tidak terelakan dari adanya persaingan antar perguruan tinggi. Hal ini juga didukung dengan adanya data yang bersumber dari lokadata.bertiatagar.id bahwa adanya pertumbuhan jumlah perguruan tinggi dari tahun ke tahun, pada 2008 tercatat terdapat 2.680 perguruan tinggi negeri dan swasta hingga pada tahun 2017 tercatat terdapat 3.276 perguruan tinggi negeri dan swasta dibawah Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi [2]. Persaingan yang ketat antar perguruan tinggi tidak luput dari masing-masing perguruan tinggi yang kerap berusaha untuk memberikan performa yang terbaik, meningkatkan kualitas dan pemberian pendidikan terbaik bagi para mahasiswanya [3]. Parameter yang umumnya dipakai dalam mengukur kualitas perguruan tinggi yaitu akreditasi. Akreditasi merupakan alat ukur yang dikelola oleh Badan Akreditasi Nasional Perguruan Tinggi yang umumnya disingkat sebagai BAN PT, salah satu aspek yang mendasari nilai akreditasi yaitu mahasiswa dan lulusan yang salah satu komponen yaitu rata-rata masa studi atau waktu kelulusan [1].

Kelulusan merupakan serangkaian tahapan yang perlu dilalui atau diselesaikan mahasiswa, seperti melakukan kerja praktik, mengambil dan menyelesaikan jumlah minimal matakuliah, seminar proposal penelitian, dan hal lainnya yang telah ditetapkan oleh pihak institusi [4]. Pada faktanya, berdasarkan Laporan Statistik Pendidikan Tinggi pada tahun 2020, terlihat bahwa sebanyak 478.826 atau 79,5% mahasiswa putus kuliah di perguruan tinggi swasta [5]. Faktor-faktor terjadinya putus kuliah pada mahasiswa sangatlah beragam dan salah satunya yaitu dikarenakan masa studi yang melewati maksimal durasi, ditambah adanya peraturan yang telah ditetapkan pemerintah terkait Standar Nasional Pendidikan

Tinggi (SNPT) yang tertuang pada Peraturan Menteri Pendidikan dan Kebudayaan yang berisikan berkurangnya durasi atau masa studi maksimal pendidikan sarjana (S-1) dari 7 tahun menjadi 5 tahun [6].

Keterlambatan waktu kelulusan hingga putus kuliah tidak hanya merugikan pihak universitas namun juga dari sisi mahasiswa yang tidak hanya mengeluarkan uang berlebih namun juga waktu dan tenaga [7]. Kelulusan yang tertunda hingga terjadinya putus kuliah menimbulkan dampak yang krusial di seluruh bagian dunia. Hal ini dapat berdampak secara signifikan bagi para individu yang mengalaminya seperti menimbulkan tekanan psikologis hingga telambatnya pencapaian karir [8]. Berdasarkan hal tersebut, bimbingan dan arahan sebagai salah satu bentuk konseling mahasiswa menjadi salah satu layanan yang vital dalam memperteguh kesuksesan akademik mahasiswa [9].

Universitas Multimedia Nusantara sebagai salah satu universitas swasta yang telah berdiri sejak tahun 2005, sudah melaksanakan kosep dengan terdapatnya dosen pembimbing akademik yang dapat membantu dan mengarahkan mahasiswa bimbingannya dalam merancang rencana studi. Pada pelaksanaannya, dosen pembimbing akademik harus menganalisa hasil studi dari waktu ke waktu untuk setiap mahasiswa bimbingannya secara manual. Berdasarkan permasalahan tersebut, penelitian ini berfokus untuk melakukan pembuatan model prediksi tipe kelulusan mahasiswa dengan menerapkan konsep *educational data mining*. Model prediksi berfungsi sebagai salah satu acuan dalam melakukan kontrol dan *tracking* perjalanan studi mahasiswa yang diharapkan menjadi salah satu tindakan preventif dari terjadinya kondisi mahasiswa yang lulus tidak tepat waktu hingga putus kuliah.

Model prediksi yang dilakukan dengan memilih performa terbaik dari perbandingan algoritma *Neural Network* (NN), *Naive Bayes* (NB), dan *Support Vector Machine* (SVM) yang telah dioptimasi menggunakan metode *Particle Swarm Optimization* (PSO). Algoritma *Neural Network* dipilih menjadi salah satu algoritma yang dibandingkan karena algoritma tersebut mendapatkan hasil akurasi diatas 90% berdasarkan penelitian terdahulu [10][11]. Berdasarkan penelitian yang telah dilakukan, algoritma *Naive Bayes* juga dipilih sebagai salah satu algoritma

yang dibandingkan karena performa akurasi mencapai lebih dari 90% [1]. Algoritma *Support Vector Machine* juga dipilih sebagai salah satu pembanding algoritma karena kemampuan performan akurasi terbaik dibanding algoritma lainnya dalam penelitian terdahulu [12][13]. Penelitian ini memiliki fokus yang berbeda yaitu pertimbangan terkait fungsionalitas model klasifikasi dalam melakukan prediksi multikelas sebanyak 7 kelas yang dapat digunakan oleh mahasiswa semester 1 hingga 14 walaupun memiliki *input shape* yang berbeda. Optimasi algoritma dilakukan dengan alasan bahwa model algoritma dengan parameter *default* memiliki sifat yang lebih rentan terhadap serangan adversarial [14], serta karena sifatnya yang *default*, model algoritma dapat menunjukkan hasil yang baik pada data tertentu, tetapi tidak pada data lainnya [15]. Di sisi lain, keperluan optimasi agar mendapatkan performa yang lebih baik juga dipertimbangkan karena kebutuhan implementasi model dan pengimplementasian yang dilakukan tidak secara langsung pada *website (back-end)*. Optimasi algoritma dengan metode PSO telah terbukti dapat meningkatkan performa algoritma pada penelitian terdahulu dimana terdapat peningkatan akurasi pada algoritma Neural Network, Naïve Bayes, dan Support Vector Machine berturut-turut sebesar 5.11%, 2.13%, dan 1.79%. Akurasi algoritma Neural Network berhasil meningkat dari 67.44% menjadi 72.55%, algoritma Naïve Bayes meningkat dari 63.79% menjadi 65.92%, dan algoritma Support Vector Machine meningkat 69.10% menjadi 70.89% [16]. Pemilihan algoritma optimasi PSO juga dilakukan karena metode tradisional untuk konfigurasi *hyperparameter* seperti *grid search* atau *random search* memiliki sifat yang *time-consuming* dan *labor-intensive*. Metode *grid search* efektif jika jumlah *hyperparameter* sedikit, namun, jika jumlah *hyperparameter* menjadi besar, waktu yang diperlukan oleh metode *grid search* akan meningkat secara eksponensial. Metode *random search* memilih *hyperparameter* optimal dengan cara pengambilan sampel acak, maka dari itu metode ini memiliki tingkat ketidakpastian tertentu karena setiap pengambilan sampel tidak memperhitungkan hasil sebelumnya, sehingga dapat menyebabkan pencarian berulang [17].

Hasil dari penelitian ini adalah informasi terkait daftar para mahasiswa yang dikelompokkan berdasarkan data performa mahasiswa dan model untuk

memprediksi waktu kelulusan mahasiswa dengan dilihat dari variabel semester, satuan kredit semester (SKS), indeks prestasi semester (IPS), indeks prestasi kumulatif (IPK), dan total masa studi. Dengan hasil dari analisis yang didapatkan, diharapkan informasi dan model dapat membantu bagi Universitas Multimedia Nusantara dan mahasiswa Universitas Multimedia Nusantara untuk melakukan penyesuaian untuk meningkatkan presentase ketepatan waktu kelulusan. Selain itu, hasil dari analisis diharapkan dapat menjadi informasi tambahan bagi pihak kampus dan mahasiswa lain dalam meningkatkan ketepatan waktu kelulusan.

1.2 Rumusan Masalah

Berdasarkan latar belakang tersebut, masalah penelitian ini dirumuskan:

- 1) Bagaimana penerapan *educational data mining* dalam pembentukan model prediksi waktu kelulusan mahasiswa menggunakan algoritma klasifikasi?
- 2) Bagaimana perbandingan hasil dan kinerja algoritma Neural Network, Naïve Bayes, dan Support Vector Machine yang dioptimasi dan tidak dioptimasi menggunakan algoritma Particle Swarm Optimization dalam memahami pola dan meminimalisir kesalahan prediksi berdasarkan ukuran akurasi, presisi, *recall*, spesifisitas, dan *f1-score*?
- 3) Bagaimana implementasi model prediksi ke dalam sistem secara *web-based*?

1.3 Batasan Masalah

Batasan masalah pada penelitian yaitu:

- 1) Variabel yang digunakan untuk penelitian ini yaitu satuan kredit semester yang diambil per semester, satuan kredit semester yang lulus per semester, satuan kredit semester yang tidak lulus per semester, total satuan kredit semester yang telah diambil per semester, indeks prestasi semester per semester, indeks prestasi kumulatif per semester dan, total masa studi dalam semester.
- 2) Sistem hanya bisa digunakan di Universitas Multimedia Nusantara karena data yang digunakan merupakan data kartu hasil studi mahasiswa Sistem Informasi Universitas Multimedia Nusantara.

- 3) Data yang digunakan merupakan data mahasiswa tahun angkatan 2014-2023.
- 4) Metode evaluasi yang digunakan yaitu akurasi, presisi, sensitivitas, spesifisitas, dan *f1-score*.
- 5) *Website* yang akan dijadikan platform untuk tahap *deployment* tidak memiliki fungsi yang komprehensif karena hanya digunakan untuk simulasi dari tahap pengimplementasian model.
- 6) Algoritma optimasi menggunakan algoritma Particle Swarm Optimization (PSO)
- 7) Kuesioner *User Acceptance Testing* untuk mahasiswa disebarakan pada 10 persen dari total mahasiswa aktif.

1.4 Tujuan dan Manfaat Penelitian

1.4.1 Tujuan Penelitian

Tujuan dari penelitian ini yaitu:

- 1) Merancang model dengan teknik *educational data mining* untuk memprediksi waktu kelulusan mahasiswa.
- 2) Menilai perbandingan hasil dan kinerja algoritma Neural Network, Naïve Bayes, dan Support Vector Machine yang dioptimasi dan tidak dioptimasi menggunakan algoritma Particle Swarm Optimization dalam memahami pola dan meminimalisir kesalahan prediksi berdasarkan ukuran akurasi, presisi, *recall*, spesifisitas, dan *f1-score*.
- 3) Melakukan penerapan hasil akhir pengolahan data kedalam secara *web-based*.

1.4.2 Manfaat Penelitian

Manfaat dari penelitian ini adalah:

- 1) Manfaat Teoritis:
 - a) Implementasi *educational data mining* dengan membandingkan algoritma Naïve Bayes, Support Vector Machine, dan Neural Network yang dioptimasi menggunakan algoritma Particle Swarm Optimization dalam memprediksi waktu kelulusan mahasiswa dalam semester

diharapkan dapat memberikan wawasan tambahan bagi penelitian terkait prediksi waktu kelulusan mahasiswa.

2) Manfaat Praktis:

- a) Memberikan *insight* untuk pihak mahasiswa dan dosen pembimbing pada Universitas Multimedia Nusantara terkait waktu kelulusan masa studi mahasiswa tersebut.
- b) Membantu pihak Biro Informasi Akademik Universitas Multimedia Nusantara dalam membuat fitur prediksi waktu kelulusan mahasiswa sebagai salah satu fitur pada *website* resmi akademik.
- c) Sebagai masukan untuk meningkatkan kesadaran mahasiswa dalam melakukan antisipasi mahasiswa yang tidak atau kurang sesuai dengan target standar waktu kelulusan.

1.5 Sistematika Penulisan

Sistematika penulisan yang digunakan dalam tugas akhir ini adalah sebagai berikut:

BAB 1 PENDAHULUAN

Pendahuluan memuat latar belakang dari permasalahan yang diulas pada penelitian yang dilakukan, batasan masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

BAB 2 LANDASAN TEORI

Bagian ini memuat teori-teori pendukung yang menjadi dasar pengetahuan yang berkaitan dengan penelitian yang dilakukan. Artikel jurnal atau buku yang berkaitan dengan topik penelitian menjadi sumber teori yang dipakai pada penelitian yang dilakukan.

BAB 3 METODOLOGI PENELITIAN

Pada bab ini memuat penguraian penjelasan dari objek yang diteliti, metode yang dipakai pada penelitian, variabel yang terdapat pada objek penelitian, teknik pengumpulan data, dan teknik analisis atau pengolahan data yang akan dilakukan.

BAB 4 ANALISIS DAN HASIL PENELITIAN

Bagian ini memuat proses pengimplementasian teknik *educational data mining* terhadap objek yang diteliti. Bab ini juga mencakup pemrosesan awal, pengolahan data, dan hasil performa model.

BAB 5

SIMPULAN DAN SARAN

Bagian ini memuat kesimpulan dan saran untuk penelitian lanjutan berdasarkan hasil dan analisa yang diperoleh.

