

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Pendidikan adalah elemen krusial dalam pengembangan individu dan masyarakat, berfungsi untuk mengoptimalkan potensi diri serta meningkatkan kualitas hidup. Sejak usia dini hingga lanjut usia, manusia sangat membutuhkan pendidikan sebagai sarana untuk meningkatkan kualitas diri dan membangun potensi secara optimal. Pendidikan tidak hanya berfungsi sebagai jalur meningkatkan kualitas diri, tetapi juga sebagai pilar utama untuk menciptakan masyarakat yang lebih maju dan sejahtera guna mendukung pembangunan suatu bangsa. Dengan pendidikan kita dapat mengembangkan potensi sumber daya manusia (SDM) secara maksimal untuk menghadapi tantangan global dan perubahan zaman yang semakin kompleks. Selain itu, pendidikan juga berperan dalam menanamkan nilai-nilai moral, keterampilan hidup, dan kemampuan berpikir kritis, yang menjadi fondasi penting dalam kehidupan bermasyarakat. Investasi dalam pendidikan berkualitas adalah langkah penting untuk menciptakan masa depan yang lebih baik bagi individu dan bangsa secara keseluruhan. Pendidikan menjadi kunci utama dalam membangun bangsa yang kompetitif di tingkat global.

Proses pendidikan memiliki durasi waktu yang telah ditetapkan oleh lembaga pendidikan sesuai dengan tingkat atau jenjang pendidikan. Apabila waktu yang ditentukan terlampaui, maka peserta didik dianggap tidak berhasil dalam menyelesaikan proses pendidikannya (Winardi dkk., 2014). Proses pendidikan tersebut dilaksanakan di seluruh tingkat dan jenjang pendidikan, mulai dari

pendidikan dasar, pendidikan menengah, hingga pendidikan tinggi, yang mencakup program wajib belajar 9 tahun. Perguruan tinggi, sebagai institusi pendidikan, memiliki tanggung jawab untuk menjalankan peran dan fungsinya dalam upaya mencapai tujuan pendidikan (Zufriah, 2022). Perguruan tinggi dirancang untuk mencetak generasi yang berdaya saing tinggi dan mendukung tercapainya tujuan pendidikan nasional. Pendidikan tinggi menjadi salah satu tahap yang membekali individu dengan kompetensi mendalam untuk menghadapi tantangan dalam dunia profesional.

Di Indonesia, perguruan tinggi tidak hanya bertanggung jawab memberikan berbagai ilmu pengetahuan dan keterampilan, tetapi juga harus berperan aktif dalam memecahkan masalah yang dihadapi masyarakat. Sebagai masyarakat ilmiah, perguruan tinggi dituntut untuk menghasilkan ilmu yang dapat membantu memecahkan masalah (*problem solver*) yang dihadapi masyarakat. Dengan demikian, ilmu yang diperoleh dapat digunakan untuk menjelaskan (*to explain*), meramalkan (*to predict*), atau peristiwa (*event*) yang terjadi di masyarakat (Yuliawati, 2012). Hal ini sejalan dengan prinsip Tri Dharma Perguruan Tinggi, yaitu penelitian, pendidikan, dan pengabdian kepada masyarakat yang merupakan landasan utama bagi seluruh perguruan tinggi di Indonesia. Ketiga pilar ini sebagai bagian dari upaya untuk mewujudkan tujuan Negara Indonesia yaitu mencerdaskan kehidupan bangsa. Dengan demikian, tujuan utama dari pendidikan tinggi tidak hanya untuk menghasilkan lulusan yang kompeten dan siap bekerja di berbagai sektor tetapi juga memiliki kemampuan berpikir kritis, kreatif, dan solutif sehingga mampu bersaing dengan bangsa-bangsa lain dalam menghadapi globalisasi.

Berdasarkan Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 49 Tahun 2014 tentang Standar Nasional Pendidikan Tinggi, khususnya Pasal 17 ayat (3), disebutkan bahwa masa studi bagi mahasiswa program diploma empat dan program sarjana adalah 4 (empat) sampai 5 (lima) tahun (Kementerian Pendidikan dan Kebudayaan Republik Indonesia, 2014). Namun, pada kenyataannya, tidak semua mahasiswa dapat menyelesaikan studi mereka tepat waktu. Beberapa mahasiswa sering kali menghadapi berbagai macam hambatan selama proses studi yang menyebabkan mereka tidak dapat menyelesaikan pendidikan tepat waktu. Keterlambatan penyelesaian studi ini merupakan situasi yang umum terjadi di berbagai perguruan tinggi dan dapat muncul dari berbagai aspek yang saling berkaitan. Dampak dari keterlambatan ini tidak hanya dirasakan pada individu mahasiswa, tetapi juga pada reputasi institusi dan efisiensi sistem pendidikan yang bersangkutan. Perguruan tinggi menghadapi tantangan untuk memastikan bahwa program-program akademik mereka efektif dalam membantu mahasiswa menyelesaikan studi tepat waktu. Dalam jangka panjang, situasi ini dapat mempengaruhi kepercayaan masyarakat terhadap efektivitas sistem pendidikan tinggi di Indonesia.

Di Universitas Pendidikan Ganesha, keterlambatan penyelesaian studi menjadi salah satu isu yang tidak asing lagi untuk didengar. Perbedaan mencolok terlihat dari jumlah mahasiswa yang dapat menyelesaikan studi tepat waktu dengan mereka yang melebihi batas waktu yang ditentukan. Salah satu permasalahan yang paling sering ditemukan pada beberapa mahasiswa adalah kendala dalam menyusun tugas akhir atau skripsi yang menjadi persyaratan utama kelulusan. Banyak mahasiswa yang menganggap bahwa skripsi merupakan tugas yang sulit, memakan

waktu lama, penuh tantangan, dan rumit, sehingga menyebabkan mahasiswa merasa takut dan malas untuk mengerjakannya (Adnyani, 2023). Hal ini diperburuk oleh rendahnya minat mahasiswa untuk melaksanakan bimbingan akademik secara rutin dan keterbukaan dalam menyampaikan permasalahan yang dihadapi kepada dosen pembimbing, membuat mahasiswa sering kali mengalami kesulitan dan kebingungan dalam menyelesaikan tugas akhir mereka. Selain itu, tekanan akademik yang tinggi dan kurangnya manajemen waktu yang baik serta kebiasaan menunda-nunda pekerjaan semakin memperparah situasi ini. Akibatnya, mahasiswa sering kali kehilangan motivasi untuk melanjutkan proses studi dengan baik.

Permasalahan keterlambatan ini semakin kompleks ketika melihat kondisi di lapangan yang beragam. Berdasarkan pengamatan dan penelitian sebelumnya, berbagai faktor baik internal maupun eksternal, dapat menyebabkan keterlambatan dalam penyelesaian studi. Keterlambatan ini seringkali disebabkan oleh berbagai hambatan yang dihadapi mahasiswa, seperti kesulitan menyeimbangkan tuntutan akademik dengan kehidupan pribadi, adaptasi terhadap lingkungan kampus, dan tekanan akademik yang tinggi. Faktor internal adalah faktor yang timbul dari dalam diri individu mahasiswa seperti kurangnya motivasi belajar, manajemen waktu yang buruk, serta permasalahan kesehatan mental. Sedangkan, faktor eksternal adalah faktor-faktor yang dapat mempengaruhi hasil belajar yang sifatnya di luar diri mahasiswa, seperti keterbatasan ekonomi, kurangnya dukungan dari keluarga atau lingkungan yang dapat memperparah situasi ini (Latief dkk., 2023). Hal ini sejalan dengan penelitian oleh Hafsah (2024) dan penelitian oleh Yola Pramika dan Linden Rosalina (2021) yang mengungkapkan bahwa penyelesaian studi pada mahasiswa

disebabkan oleh dua faktor utama yaitu faktor eksternal dan faktor internal (Hafsah, 2024; Pramika & Rosalina, 2021).

Dalam upaya memahami dan mengatasi keterlambatan studi, berbagai pendekatan telah dilakukan, salah satunya melalui pemanfaatan data mining. Salah satu penelitian yang relevan dilakukan oleh Endang Etriyanti, Dedy Syamsuar, dan Yesi Novaria Kunang (2020) yang menggunakan algoritma Naïve Bayes Classifier untuk memprediksi kelulusan mahasiswa. Dalam penelitian ini peneliti memanfaatkan atribut seperti jenis kelamin, asal sekolah, IP semester 1-4, IPK semester 4 dan status kelulusan untuk memprediksi. Adapun hasil yang diperoleh adalah dengan metode Naive Bayes Classifier memperoleh nilai Accuracy 78,46% dan prediksi menggunakan metode Algoritme C4.5 diperoleh nilai Accuracy yang lebih besar yaitu 79,08%. Penelitian ini juga menunjukkan kriteria yang berpengaruh dalam prediksi tersebut yaitu IPK semester 4, jenis kelamin, dan asal sekolah (Etriyanti dkk., 2020).

Penelitian serupa juga mengungkapkan temuan menarik tentang prediksi kelulusan mahasiswa tepat. Misalnya, penelitian oleh Lila Setiyani, Mokhammad Wahidin, Dudi Awaludin, dan Sri Purwani (2020) juga mengkaji terkait prediksi yang sejalan dengan keterlambatan penyelesaian studi dan menerapkan data mining. Dalam penelitian tersebut, para peneliti memanfaatkan atribut-atribut dari database perguruan tinggi untuk memprediksi kelulusan mahasiswa tepat waktu. Hasil penelitian ini mengungkapkan bahwa metode Naïve Bayes mampu memberikan prediksi dengan tingkat akurasi lebih dari 90%, meskipun menggunakan sejumlah atribut yang berbeda. Salah satu atribut yang berulang kali ditemukan sebagai faktor prediktor utama adalah IPK (Indeks Prestasi Kumulatif),

yang memiliki peran penting dalam menentukan kelulusan mahasiswa tepat waktu. Penelitian ini menunjukkan bahwa Indeks Prestasi Kumulatif (IPK) merupakan indikator yang sangat signifikan, dan pemanfaatan data mining dalam bentuk Naïve Bayes dapat memberikan insight yang berguna bagi pihak universitas untuk mengidentifikasi mahasiswa yang berisiko mengalami keterlambatan studi (Setiyani dkk., 2020). Penelitian oleh Rafika Syahrani (2022) juga menggunakan pendekatan data mining dengan metode Regresi Logistik Multinomial untuk memprediksi kategori kelulusan mahasiswa berdasarkan variabel seperti jenis kelamin, jalur masuk, dan IPK dari semester 1 hingga 6. Hasil penelitian ini menunjukkan bahwa model Regresi Logistik Multinomial memiliki akurasi sebesar 85,5%, dengan precision 78,5%, recall 93,9%, dan micro F1-score 89,8% (Syahrani & Zaman, 2023). Kedua penelitian ini menguatkan bahwa IPK memiliki peran dominan sebagai faktor utama dalam menentukan kelulusan mahasiswa, dan terbukti memiliki pengaruh yang signifikan terhadap kelulusan tepat waktu mahasiswa, yang semakin memperkuat peranannya sebagai indikator utama dalam memprediksi keberhasilan mahasiswa menyelesaikan studi tepat waktu.

Selain itu, penelitian oleh Isra Hayati (2021) yang membandingkan algoritma Decision Tree C4.5 dan Naive Bayes dalam mengklasifikasikan mahasiswa berpotensi drop out juga relevan untuk memperkaya pemahaman tentang keterlambatan penyelesaian studi. Hasil penelitian ini menunjukkan bahwa algoritma Decision Tree C4.5 memiliki performa yang lebih tinggi dalam mengklasifikasikan mahasiswa berpotensi drop out, dengan nilai akurasi mencapai 96,74%. Pada penelitian ini, atribut SKS Total ditemukan sebagai faktor paling

berpengaruh dalam klasifikasi mahasiswa yang berpotensi drop out, yang menunjukkan bahwa jumlah SKS yang telah ditempuh mahasiswa merupakan indikator penting dalam memprediksi kelulusan tepat waktu atau keterlambatan studi. Penemuan ini memberikan bukti lebih lanjut bahwa atribut akademik, seperti SKS, memiliki peran krusial dalam menentukan keberhasilan penyelesaian studi tepat waktu (Hayati, 2021a).

Penelitian oleh Desy Iba Ricoida, Dedy Hermanto, Rusbandi, Desi Pibriana, dan M. Rizky Pribadi (2020) juga memberikan kontribusi signifikan dalam memahami keterlambatan studi mahasiswa. Dalam penelitian ini, berbagai algoritma diuji untuk memprediksi keterlambatan studi mahasiswa, dengan algoritma Random Forest menunjukkan tingkat akurasi yang paling tinggi, yaitu sebesar 90,88%. Nilai AUC yang diperoleh sebesar 97,2%, dan perhitungan F1-Score mencapai 89,9%, yang merupakan nilai tertinggi dibandingkan algoritma lainnya. Sementara itu, algoritma Decision Tree dan Logistic Regression memiliki nilai akurasi masing-masing sebesar 89,12% dan 89,47%. Walaupun akurasi Logistic Regression sedikit lebih tinggi, nilai F1-Score Decision Tree lebih baik, yaitu 88,7% dibandingkan 87,6% untuk Logistic Regression. Perbandingan ini menunjukkan bahwa meskipun berbagai algoritma memiliki tingkat akurasi yang berbeda, masing-masing memiliki kekuatan dan kontribusinya dalam membantu memprediksi keterlambatan studi mahasiswa (Ricoida dkk., 2024a).

Dengan melihat beberapa hasil penelitian sebelumnya, keterlambatan penyelesaian studi merupakan masalah yang kompleks yang melibatkan berbagai atribut. Namun, penelitian terdahulu memiliki beberapa keterbatasan yang perlu diperbaiki. Salah satu keterbatasan utama dalam penelitian terdahulu adalah

kurangnya fokus pada mahasiswa yang telah mencapai semester ke-10 atau lebih. Metode yang digunakan sebagian besar lebih berfokus pada algoritma Naïve Bayes dan Decision Tree (C4.5), sementara metode seperti *Logistic Regression* dan lainnya masih jarang dieksplorasi dalam konteks prediksi keterlambatan studi. Beberapa atribut yang menonjol seperti IPK dan SKS belum dimanfaatkan dengan baik. Padahal, IPK semester keempat dan kedelapan mencerminkan profil keberhasilan belajar mahasiswa yang dapat digunakan secara optimal oleh dosen dalam membimbing mahasiswa yang bersangkutan. IPK sendiri merupakan indikator penting untuk menilai sejauh mana penguasaan kompetensi mahasiswa terhadap mata kuliah yang telah ditempuh, sementara SKS menunjukkan kemajuan secara kuantitatif. Namun demikian, data IPK semester delapan sering sekali belum tersedia untuk mahasiswa aktif terutama yang belum mencapai semester akhir.

Untuk mengatasi keterbatasan yang ada, dengan kemajuan teknologi informasi, data mining menjadi alat yang efektif untuk menganalisis data-data besar dan menemukan pola yang tersembunyi yang dapat membantu memahami atribut-atribut yang memengaruhi keterlambatan studi. Algoritma yang digunakan adalah regresi linier dan regresi logistik, yang diterapkan secara bertahap. Regresi linier merupakan metode statistik yang digunakan untuk menganalisis hubungan antara satu variabel dependen dengan satu atau lebih variabel independen dalam bentuk garis lurus serta bersifat kontinu atau numerik. Sementara itu, regresi logistik adalah metode statistika yang digunakan untuk mengidentifikasi hubungan antara satu atau beberapa variabel independen dengan variabel dependen yang bersifat kategorik (Hosmer & Lemeshow, 2000). Algoritma ini tidak hanya mampu mengidentifikasi faktor-faktor yang signifikan, tetapi juga dapat memberikan

estimasi probabilitas keterlambatan berdasarkan faktor-faktor tersebut serta mengevaluasi akurasi model prediksi dalam mengidentifikasi mahasiswa yang berisiko terlambat menyelesaikan studi.

Berdasarkan pemaparan di atas, dilakukan penelitian pengembangan dengan judul dengan judul **“Penerapan Data Mining Untuk Memprediksi Keterlambatan Penyelesaian Studi Mahasiswa S1 Universitas Pendidikan Ganesha”**.

1.2 Identifikasi Masalah

Berdasarkan latar belakang yang telah dipaparkan, maka peneliti memperoleh beberapa masalah diidentifikasi yakni sebagai berikut:

1. Masih banyak mahasiswa yang mengalami keterlambatan dalam menyelesaikan studi, melebihi masa studi ideal yang telah ditetapkan oleh lembaga pendidikan.
2. Mahasiswa sering mengalami kesulitan dalam menyusun tugas akhir atau skripsi, yang merupakan syarat utama kelulusan.
3. Terdapat berbagai faktor akademik dan non-akademik yang menyebabkan keterlambatan penyelesaian studi mahasiswa, namun belum teridentifikasi dengan jelas.
4. Pemanfaatan teknologi seperti data mining dalam analisis data akademik mahasiswa untuk memprediksi keterlambatan studi masih belum optimal.
5. Diperlukan penerapan data mining untuk membantu memahami faktor-faktor yang mempengaruhi keterlambatan penyelesaian studi, serta untuk memprediksi mahasiswa yang berisiko mengalami keterlambatan.

1.3 Pembatasan Masalah

Agar pembahasan dalam penelitian ini lebih terarah dan fokus pada tujuan utama, maka ditetapkan beberapa batasan masalah sebagai berikut.

1. Penelitian ini hanya mencakup mahasiswa aktif dan mahasiswa yang telah dinyatakan lulus dari program Sarjana (S1) Universitas Pendidikan Ganesha.
2. Penelitian ini tidak mengelompokkan mahasiswa berdasarkan program studi, sehingga analisis dilakukan secara keseluruhan terhadap mahasiswa S1 Universitas Pendidikan Ganesha.
3. Penelitian difokuskan untuk memprediksi keterlambatan penyelesaian studi yang dalam penelitian ini didefinisikan sebagai mahasiswa yang menempuh masa studi lebih dari 10 semester atau 5 tahun.
4. Analisis diarahkan pada mahasiswa semester 8 ke atas, karena tahap ini mahasiswa dianggap sudah mendekati akhir masa studi dan dapat diprediksi kemungkinan keterlambatannya.
5. Atribut yang dianalisis dibatasi pada atribut akademik yang tersedia dalam data, seperti jalur masuk, sumber biaya, Uang Kuliah Tunggal (UKT), Satuan Kredit Semester, serta Indeks Prestasi Kumulatif (IPK) semester 8, 6, dan 8.

1.4 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan, maka munculah sebuah rumusan masalah yakni sebagai berikut:

1. Bagaimana hasil penerapan algoritma regresi linier dalam memprediksi IPK semester 8?
2. Bagaimana hasil penerapan algoritma regresi logistik dalam memprediksi

keterlambatan penyelesaian studi mahasiswa S1 Universitas Pendidikan Ganesha?

3. Seberapa akurat model regresi logistik dalam memprediksi keterlambatan penyelesaian studi mahasiswa S1 Universitas Pendidikan Ganesha?
4. Apa saja atribut yang paling berpengaruh dalam memprediksi keterlambatan penyelesaian studi mahasiswa S1 Universitas Pendidikan Ganesha?

1.5 Tujuan Penelitian

Berdasarkan latar belakang yang telah dipaparkan, adapun tujuan penelitian ini sebagai berikut:

1. Untuk mengetahui hasil penerapan algoritma regresi linier dalam memprediksi IPK semester 8.
2. Untuk mengetahui hasil penerapan algoritma regresi logistik dalam memprediksi keterlambatan penyelesaian studi mahasiswa S1 Universitas Pendidikan Ganesha.
3. Untuk mengetahui akurasi model regresi logistik dalam memprediksi keterlambatan studi mahasiswa S1 Universitas Pendidikan Ganesha.
4. Untuk mengetahui atribut-atribut yang berpengaruh dalam memprediksi keterlambatan studi mahasiswa S1 Universitas Pendidikan Ganesha.

1.6 Manfaat Penelitian

Dengan adanya penelitian ini, diharapkan dapat memberikan manfaat teoritis dan manfaat praktis sebagai berikut:

1. Manfaat Teoritis

Penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan

ilmu pengetahuan, khususnya dalam bidang data mining yang diterapkan diarahkan Pendidikan tinggi. Melalui penerapan algoritma tersebut, penelitian ini tidak hanya memperkaya literatur mengenai teknik memprediksi akademik tetapi juga memberikan gambaran bagaimana data akademik mahasiswa dapat dianalisis secara kuantitatif untuk mendukung langkah yang akan diambil. Selain itu, penelitian ini juga membantu menjembatani pemahaman teoritis dengan praktik nyata, khususnya dalam menerapkan teknik prediksi.

2. Manfaat Praktis

a. Bagi Penulis

Diharapkan penelitian ini dapat meningkatkan keterampilan penulis dalam melakukan pengolahan dan analisis data dengan menggunakan data mining, khususnya algoritma regresi logistik. Penulis memperoleh pemahaman mendalam mengenai penerapan teknik ini, termasuk interpretasi hasil analisis. Selain itu, penelitian ini memungkinkan penulis untuk berkontribusi secara langsung dalam upaya mengatasi keterlambatan mahasiswa dalam menyelesaikan studinya, khususnya di Universitas Pendidikan Ganesha.

b. Bagi pembaca

Diharapkan penelitian ini dapat memberikan gambaran yang jelas dalam memprediksi keterlambatan penyelesaian studi mahasiswa, termasuk memanfaatkan prediksi IPK semester 8 sebagai salah satu indikator pendukung serta atribut mana yang berpengaruh dalam keterlambatan penyelesaian studi mahasiswa S1. Informasi yang dihasilkan dapat

menjadi dasar yang digunakan oleh pihak universitas, khususnya pimpinan, dosen, dan tenaga pendidik, dalam menyusun kebijakan atau strategi untuk meningkatkan keberhasilan penyelesaian studi mahasiswa. Selain itu, pembaca juga mendapatkan wawasan mengenai bagaimana penerapan data mining, khususnya regresi logistik dapat digunakan dalam permasalahan sosial, sehingga dapat diaplikasikan pada studi kasus lain yang serupa.

