

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perguruan tinggi memiliki peran strategis dalam menciptakan sumber daya manusia yang berkualitas melalui proses pendidikan yang efektif dan efisien. Setiap mahasiswa diharapkan dapat menyelesaikan studinya tepat waktu sesuai dengan kurikulum dan standar akademik yang telah ditetapkan. Namun, kenyataan menunjukkan bahwa tidak sedikit mahasiswa yang mengalami kendala sehingga gagal menyelesaikan studi tepat waktu dan akhirnya mengalami putus studi (*drop out*). Fenomena *drop out* ini berdampak negatif terhadap kemajuan perguruan tinggi, termasuk menurunkan kualitas lulusan dan reputasi perguruan tinggi. Karena dampak mahasiswa *drop out* tersebut pastinya akan menghambat kemajuan perguruan tinggi.

Politeknik Negeri Bengkalis sebagai salah satu perguruan tinggi vokasi di Indonesia yang berfokus pada pendidikan terapan menghadapi tantangan serupa. Karakteristik mahasiswa yang berbeda antar jurusan memengaruhi prestasi akademik dan kemampuan menyelesaikan studi tepat waktu. Beberapa jurusan di Politeknik Negeri Bengkalis menghadapi masalah tingginya potensi *drop out* yang dapat menghambat efektivitas proses pendidikan dan menurunkan mutu lulusan. Berdasarkan data, sejak tahun 2023 hingga 2024 telah tercatat sebanyak 198 mahasiswa yang dinyatakan *drop out* melalui Surat Keputusan (SK) di Politeknik Negeri Bengkalis. Oleh karena itu, identifikasi dini mahasiswa yang berpotensi *drop out* menjadi langkah penting agar perguruan tinggi dapat memberikan tindakan yang tepat, seperti bimbingan akademik maupun konseling untuk meningkatkan keberhasilan studi mahasiswa.

Sejumlah penelitian terdahulu telah melakukan klasifikasi terhadap mahasiswa berpotensi *drop out* menggunakan *machine learning*. Salah satu yang berkaitan yaitu penelitian oleh [1] yaitu klasifikasi mahasiswa berpotensi *drop out* menggunakan metode *K-Nearest Neighbor* dan *Random Forest* dimana menggunakan 6 variabel penelitian diantaranya IPK, total SKS, penghasilan orang tua, jalur bayaran, status bayaran dan presentase aktif. Variabel yang paling berpengaruh yaitu presentase aktif. Dan menghasilkan akurasi tertinggi oleh *Random Forest* dengan hasil akurasi 99.05% dibanding *K-Nearest Neighbor* dengan akurasi 98.10%.

Penelitian lainnya oleh [2] mengenai prediksi mahasiswa *drop out* yang memanfaatkan pendekatan *machine learning* dalam bentuk prediksi. Penelitian ini telah diupayakan supaya mampu mendeteksi tingkat *drop out* mahasiswa berdasarkan informasi historis akademik dan non-akademik. Beberapa variabel penting yang dimanfaatkan pada penelitian ini seperti IPK, jumlah SKS, kehadiran, status pembayaran, serta informasi latar belakang sosial ekonomi. Variabel yang berpengaruh yaitu IPK dan kehadiran. Algoritma yang digunakan adalah Regresi Logistik dan *Random Forest*. *Random Forest* menghasilkan tingkat akurasi untuk prediksi sebesar 95.8% dan Regresi Logistik dengan akurasi 91.3%.

Penelitian lain oleh [3] menggunakan algoritma *Naive Bayes* pada klasifikasi mahasiswa berpotensi putus studi di Fakultas Teknik Unimush Makassar mendapatkan akurasi sebesar 97%. Penelitian oleh [4] menerapkan metode *Support Vector Machine* pada klasifikasi *drop out* mahasiswa di Politeknik Negeri Lhokseumawe dengan akurasi 86%. Serta penelitian lainnya oleh [5] menggunakan algoritma *Decision Tree C4.5* pada klasifikasi mahasiswa *drop out* di Universitas Advent Indonesia mendapatkan akurasi 90%.

Klasifikasi sangat diperlukan dalam penelitian mengklasifikasikan mahasiswa berpotensi *drop out* karena dapat membantu perguruan tinggi untuk deteksi dini mahasiswa yang beresiko tinggi mengalami *drop out* sehingga bimbingan akademik, konseling atau bantuan keuangan dapat dilakukan lebih awal. Sehingga dengan pendekatan berbasis data, pengambilan keputusan menjadi lebih

tepat sasaran dan efisien terutama dalam memantau banyak mahasiswa tanpa harus mengevaluasi satu per satu secara manual. Pengklasifikasian mampu memetakan faktor-faktor yang paling berpengaruh terhadap resiko *drop out* seperti IPK, kehadiran dan bahkan latar belakang sosial yang sangat berguna dalam penyusunan kebijakan kampus. Dengan menurunnya angka *drop out* melalui klasifikasi dan penanganan yang tepat, kualitas pendidikan dan reputasi perguruan tinggi juga akan meningkat. Sehingga tindakan yang di ambil selalu memberikan solusi yang paling efektif dalam upaya pencegahan mahasiswa *drop out*.

Penggunaan algoritma *Random Forest* dalam penelitian ini karena dilihat pada keunggulannya dalam menangani data dengan banyak atribut dan kemampuannya memastikan bahwa model tidak *overfitting*. *Random Forest* bekerja dengan membangun banyak pohon keputusan (*Decision Tree*) dan menentukan hasil klasifikasi berdasarkan suara mayoritas sehingga menghasilkan model yang akurat dan stabil.

1.2 Rumusan Masalah

Rumusan masalah pada penelitian ini yaitu bagaimana membangun sistem klasifikasi mahasiswa Politeknik Negeri Bengkalis yang berpotensi mengalami *drop out* menggunakan algoritma *Random Forest* dan bagaimana tingkat akurasi, presisi dan recall dari algoritma *Random Forest* dalam mengklasifikasikan mahasiswa berpotensi *drop out*?

Batasan masalah pada penelitian ini yaitu data yang digunakan merupakan data mahasiswa dari Politeknik Negeri Bengkalis.

1.3 Tujuan

Tujuan penelitian ini adalah untuk membangun sistem klasifikasi mahasiswa Politeknik Negeri Bengkalis yang berpotensi *drop out* menggunakan algoritma *random forest* serta mengevaluasi kinerja algoritma *random forest* dalam mengklasifikasi mahasiswa berpotensi *drop out* dengan mengukur tingkat akurasi, presisi dan *recall*.

1.4 Manfaat

1. Meningkatkan deteksi dini mahasiswa yang berpotensi *drop out* dengan *machine learning*.
2. Mendukung pengambilan keputusan berbasis data berdasarkan hasil klasifikasi dan identifikasi pada faktor yang mempengaruhi.
3. Penurunan angka *drop out* sehingga dapat berkontribusi pada peningkatan kualitas pendidikan yang berdampak positif pada reputasi dan akreditasi Politeknik Negeri Bengkalis.
4. Penurunan angka *drop out* yang dapat meningkatkan jumlah lulusan yang lebih banyak, berkualitas dan bermanfaat.