

# BAB 1

## PENDAHULUAN

Bab Pendahuluan ini merupakan bagian awal penelitian skripsi, yang membahas latar belakang permasalahan yang terjadi, merumuskan permasalahan dari latar belakang, membatasi permasalahan, tujuan penelitian, serta manfaat penelitian tersebut.

### 1.1 Latar Belakang Masalah

Pada setiap tahunnya universitas selalu melepaskan mahasiswanya yang telah lulus dan bergelar. Jika dilihat dari informasi kelulusan mahasiswa, mahasiswa yang lulus maupun belum lulus terdapat peningkatan atau penyusutan setiap tahunnya. Banyak mahasiswa yang lulus tepat pada waktunya, namun tidak jarang pula mahasiswa lulus tidak tepat pada waktunya.

Data mahasiswa di suatu jurusan atau fakultas sangatlah melimpah, melimpahnya data pada Sistem Akademik Perguruan Tinggi seperti data jumlah kelulusan setiap tahun, hasil akademik mahasiswa, maupun informasi tersembunyi lainnya bisa dimanfaatkan dengan baik dengan melakukan pengolahan agar menjadi informasi yang berguna bagi universitas (Tambunan, Ratu H. 2020). Hal ini merupakan suatu kesempatan yang bagus untuk menganalisis studi mahasiswa selama belajar di universitas.

Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jakarta (FIK UPN Veteran Jakarta) merangkum data mahasiswanya dan mengolahnya sesuai dengan apa yang dibutuhkan. Peningkatan jumlah data diakibatkan dari tidak seimbang nya mahasiswa yang masuk dan yang sudah lulus membuat penumpukan dan jumlah mahasiswa yang semakin banyak. Data tersebut haruslah diolah dengan baik dengan teknik yang tepat.

Maka dari itu, untuk menganalisis dan mengklasifikasikan mahasiswa yang lulus digunakanlah teknik *data mining* untuk mengklasifikasikan ketepatan lama studi mahasiswa.

Dari penjelasan diatas, dilakukanlah analisis klasifikasi ketepatan kelulusan mahasiswa dengan menggunakan algoritma *Random Forest* dan *Gradient Boosting* untuk mengetahui akurasi yang terbaik. Analisis ini akan diterapkan pada FIK UPN Veteran Jakarta untuk menjadi informasi yang berguna dan untuk mengetahui mahasiswa yang memiliki potensi untuk lulus tepat pada waktunya ataupun lulus terlambat.

## 1.2 Rumusan Masalah

Dari latar belakang yang sudah dijabarkan, maka dari itu rumusan masalah bisa dirangkum sebagai berikut.

- a. Algoritma mana yang memiliki tingkat akurasi terbaik antara Random Forest dan Gradient Boosting dalam mengklasifikasikan ketepatan lulus mahasiswa?
- b. Bagaimana algoritma Random Forest dan Gradient Boosting mampu mengklasifikasikan kelulusan mahasiswa?

## 1.3 Batasan Masalah

- a. Data mahasiswa S1 FIK UPN Veteran Jakarta angkatan tahun 2015-2017.
- b. Data mahasiswa yang diklasifikasi hanya S1 Informatika dan S1 Sistem Informasi.
- c. Metode yang digunakan adalah klasifikasi dengan menggunakan Algoritma Random Forest dan Gradient Boosting.
- d. Menggunakan parameter atribut Program Studi, Jenis Kelamin, Jenis Sekolah, Provinsi Sekolah, Jalur Masuk, Beasiswa/ Non, Tarif UKT, IPK, IP Semester 1 – 7.
- e. Algoritma terbaik dilihat dari analisis performa yang terdiri dari *accuracy*, *precision*, dan *recall*.
- f. Hasil dalam bentuk daftar klasifikasi ketepatan lulus mahasiswa.

## 1.4 Tujuan Penelitian

- a. Mengetahui perbandingan akurasi dari algoritma Random Forest dan Gradient Boosting dalam mengklasifikasikan ketepatan lama studi mahasiswa.
- b. Memberikan informasi klasifikasi mahasiswa yang lulus maupun tidak lulus tepat waktu.
- c. Membantu fakultas dalam mengevaluasi kelulusan mahasiswanya pada sistem akademik fakultas yang berjalan.

## 1.5 Manfaat Penelitian

- a. Pihak pengguna dapat mengetahui tingkat keakuratan klasifikasi ketepatan lulus mahasiswa dengan menggunakan algoritma Random Forest dan Gradient Boosting.
- b. Memberikan rekomendasi penelitian ini ke fakultas untuk mempermudah dalam memperoleh informasi mahasiswa yang tepat dan akurat dalam menyelesaikan masa studinya secara tepat waktu serta meminimalisir mahasiswa yang lulus tidak tepat waktu.

## 1.6 Sistematika Penulisan

Sistematika Penulisan penelitian ini terbagi menjadi bab yang terstruktur untuk mempermudah pembaca memahaminya. Pada tiap bab mempunyai sub judul yang berguna untuk merinci bab sebelumnya. Berikut gambaran teknis penulisan untuk penelitian ini.

### BAB I PENDAHULUAN

Bab ini memiliki latar belakang masalah, rumusan masalah, identifikasi masalah, tujuan hingga manfaat penelitian.

### BAB II LANDASAN TEORI

Bab ini berisi teori yang berhubungan dengan penelitian ini. Bab ini juga menjelaskan mengenai *data mining* dan algoritma khususnya Random Forest dan Gradient Boosting. Bacaan yang dipakai berasal dari internet maupun dokumentasi buku referensi.

### BAB III METODOLOGI PENELITIAN

Bab ini berisi tahapan dan kerangka kerja penelitian. Mulai dari teknik pengumpulan data, bahan dan alat, variabel penelitian yang digunakan, jadwal penelitian dan tahapan data mining serta algoritma yang digunakan.

### BAB IV HASIL DAN PEMBAHASAN

Bab ini merinci tahapan data mining dan algoritma untuk menunjukkan hasil dan analisis penelitian ini serta bagaimana menjelaskan proses analisa dari awal sampai akhir yang diharapkan.

### BAB V PENUTUP

Bab terakhir berisi kesimpulan dan juga saran untuk pengembangan penelitian tahap selanjutnya.

### DAFTAR PUSTAKA

### RIWAYAT HIDUP

### LAMPIRAN – LAMPIRAN