

# BAB 1 PENDAHULUAN

## 1.1 Latar Belakang

Institut Teknologi Telkom Surabaya atau biasa dikenal dengan ITTelkom Surabaya adalah salah satu perguruan tinggi IT yang berada di kota Surabaya, ITTelkom Surabaya sendiri merupakan salah satu perguruan tinggi dibawah naungan Telkom Group khususnya bagian Yayasan Pendidikan Telkom. ITTelkom Surabaya didirikan pada 4 September 2018 di Jalan Ketintang No. 156, Surabaya, Jawa Timur, Indonesia. ITTelkom Surabaya hingga saat ini memiliki 11 program studi yang dibagi menjadi 2 fakultas yang berbeda, yaitu: Fakultas Teknologi Elektro dan Industri Cerdas dan Fakultas Teknologi Informasi dan Bisnis [1]. ITTelkom Surabaya memiliki akreditasi Baik sejak 2020. Aspek kelulusan pada sebuah universitas ataupun institusi merupakan salah satu aspek penting dalam peninjauan akreditasi [2]. Jika pada sebuah kampus terdapat banyak mahasiswanya lulus dengan tenggat waktu yang melebihi waktu seharusnya maka itu akan berpengaruh dalam penilaian akreditasi kampus tersebut [3].

Studi Sarjana sendiri memiliki tenggat waktu umumnya 4 tahun dengan 8 semester atau mahasiswa dapat melakukan lulus cepat dengan waktu 3,5 tahun dan 7 semester untuk mencapai kelulusan, jika mahasiswa lulus dalam tenggat waktu tersebut maka mereka akan dibilang lulus tepat waktu namun jika mereka melebihi waktu tersebut maka mereka akan dinyatakan lulus tidak tepat waktu [4]. Banyak keuntungan apabila seorang mahasiswa lulus dengan tepat waktu, salah satunya adalah jika mahasiswa lulus tepat waktu maka mereka dapat melakukan pekerjaan tanpa adanya keterikatan dengan dunia pendidikan khususnya sarjana atau kampus [2]. Namun yang jadi permasalahan disini adalah waktu kelulusan setiap mahasiswa tidak menentu, dimana ini dapat dipengaruhi dengan banyak hal mulai dari aspek *internal* mahasiswa tersebut hingga aspek-aspek *eksternal*-nya seperti lingkungannya. Maka dari itu pentingnya untuk mengetahui waktu kelulusan seorang mahasiswa meskipun hanya prediksi semata namun hal tersebut tentunya dapat membantu kampus khususnya ITTelkom Surabaya dalam mendidik mahasiswanya agar dapat lulus tepat waktu sesuai dengan tingkatan yang diambil yaitu sarjana dengan waktu 3,5 hingga 4 tahun.

Banyak cara untuk dapat memprediksi kelulusan mahasiswa dan salah satunya yang sering digunakan adalah *Data Mining*, karena *data mining* akan mencari informasi tersembunyi dari tumpukan data yang ada yang dirasa tidak memiliki fungsi lainnya [5, 6]. Pada *data mining* sendiri memiliki alur yang harus dipenuhi sebelum masuk kedalam pengolahan data dengan atau tidak dengan bantuan aplikasi pemrosesan, salah satu alur yang harus dipenuhi yaitu menentukan algoritma yang akan digunakan karena dengan menggunakan algoritma yang tepat pada data yang tepat maka akan menghasilkan informasi yang lebih akurat dan tepat. Algoritma yang dapat membantu dalam melakukan prediksi kelulusan adalah: *Decision Tree ID3*, *Decision Tree C4.5*, *Artificial Neural Networks (ANN)*, *K-Nearest Neighbour (KNN)*, *Naïve Bayes*, dan beberapa algoritma lainnya [2]. Setiap algoritma tersebut memiliki fungsi dan cara yang berbeda-beda untuk mengolah datanya demi menghasilkan informasi yang diinginkan, seperti contoh pada algoritma ID3 dan C4.5 haruslah memiliki atribut *label* untuk dijadikan sebagai hasil *output* dari algoritma tersebut dan tiap-tiap algoritma memiliki kententuannya masing-masing yang haruslah dipenuhi oleh pengguna.

Algoritma yang digunakan pada penelitian ini memiliki fokus pada algoritma *Decision Tree C4.5* karena metode ini biasanya digunakan untuk dapat mengklasifikasikan data dengan atribut numerik. Sudah banyak penelitian lainnya yang serupa untuk mencari prediksi kelulusan dengan menggunakan algoritma yang sama, salah satunya adalah yang dilakukan oleh Lydia Yohana Lumban Gaol, M. Safii, dan Dedi Suhendro pada tahun 2021 dengan menggunakan metode yang sama memberikan akurasi sebesar 90,00% dengan tingkat presisi 91,38% dan tingkat *recall* 98,15%, adapun juga penelitian lainnya yaitu oleh Muhammah Kamil dan Widya Cholil pada tahun 2020 dengan menggunakan perbandingan metode algoritma C4.5 dengan *Naïve Bayes* dan hasilnya algoritma C4.5 memiliki tingkat *confusion matrix* yang lebih tinggi dari *naïve bayes*, dan yang terakhir merupakan penelitian oleh Fuad Nurhasan, Noer Hikmah, dan Dwi Yuni Utami pada tahun 2018 dengan membandingkan 3 metode yaitu C4.5, *Naïve bayes*, dan juga KNN dengan hasil akurasi tertinggi adalah *Naïve bayes*. Dengan menggunakan aplikasi RapidMiner untuk melakukan perhitungan dapat menghasilkan pohon keputusan (*Decision Tree*) dan juga hasil penentuan *label* untuk perhitungan *confusion matrix*.

Pada *confusion matrix* sendiri memiliki 3 hasil utama, yaitu: akurasi, *recall*, dan presisi yang tersedia dengan bantuan operator *performance* pada RapidMiner.

Data yang akan digunakan pada penelitian ini adalah data yang berasal dari ITTelkom Surabaya, data angkatan 2018 akan menjadi data *training* dan data *testing* dimana data angkatan 2020 akan menjadi data perhitungan prediksi. Dari perhitungan data *training* dan data *testing* akan menghasilkan *rule* yang akan menjadi landasan dalam perhitungan prediksi kelulusan angkatan 2020 di ITTelkom Surabaya. Hasil tersebut dapat digunakan pihak kampus untuk menjadi acuan dalam mengatur strategi terhadap angkatan 2020 dan yang terbaru khususnya dalam mengantisipasi adanya keterlambatan kelulusan dimasa yang mendatang. Dari penjelasan-penjelasan sebelumnya dapat ditarik kesimpulan bahwa waktu kelulusan seorang mahasiswa dapat mempengaruhi akreditasi sebuah universitas atau insitusi dimana hal tersebut dapat diantisipasi dengan melakukan prediksi kelulusan mahasiswa yang dirasa akan terlambat dalam waktu kelulusan khususnya pada program Sarjana di ITTelkom Surabaya dengan menggunakan *Data Mining* dan algoritma *Decision Tree C4.5*, untuk aplikasi pembantu dalam menghitung *Data Mining* adalah RapidMiner.

## 1.2 Rumusan Masalah

Dengan latar belakang tersebut, permasalahan yang akan diselesaikan oleh peneliti yaitu:

1. Bagaimana cara untuk memprediksi kelulusan mahasiswa ITTelkom Surabaya dengan menggunakan algoritma *Decision Tree C4.5*?
2. Seberapa besar tingkat akurasi yang didapat dari hasil metode data mining menggunakan algoritma *Decision Tree C4.5* untuk memprediksi kelulusan mahasiswa?

## 1.3 Tujuan dan Manfaat

Dengan rumusan masalah yang ada, tujuan yang ingin dicapai oleh peneliti didalam penelitian ini adalah:

1. Mengetahui ketepatan waktu kelulusan mahasiswa ITTelkom Surabaya khususnya angkatan 2020.
2. Mengetahui tingkat akurasi pada prediksi mahasiswa ITTelkom Surabaya.

Manfaat yang akan didapatkan dari penelitian ini akan menjadi bantuan bagi pihak kampus maupun mahasiswa kedepannya, yaitu:

1. Manfaat bagi kampus adalah sebagai acuan data pada angkatan mahasiswa tersebut yang dirasa akan terlambat dalam kelulusan program sarjana untuk dapat dilakukan peninjauan lebih terhadap angkatan tersebut.
2. Manfaat bagi mahasiswa adalah sebagai sarana dalam memperbaiki diri untuk menjadi lebih giat agar dapat mengejar waktu kelulusan tepat waktu dengan waktu kelulusan program sarjana.

#### **1.4 Batasan Masalah**

Pada penelitian yang akan dikerjakan terdapat beberapa batasan-batasan yang sudah ditentukan oleh peneliti dalam melakukan penelitian agar memiliki batas dalam mengerjakan penelitian, yaitu:

1. Dilakukan pada Institut Teknologi Telkom Surabaya.
2. Penggunaan Algoritma Decision Tree C4.5 sebagai penghitung utama.
3. Data Training dan testing yang digunakan dari angkatan 2018 dan data angkatan 2020 akan dijadikan pengujian hasil rule untuk memprediksi kelulusan mahasiswa.