

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Salah satu indikator keberhasilan perguruan tinggi dalam proses belajar mengajar adalah angka kelulusan. Tingginya angka kelulusan dapat dianggap sebagai sebuah prestasi bagi sebuah perguruan tinggi. Perguruan tinggi yang memiliki lulusan-lulusan dengan daya saing yang tinggi tentunya memiliki mutu yang baik dalam hal ketepatan waktu lulus. Mahasiswa dikatakan lulus tepat waktu apabila dapat menyelesaikan masa studi dalam rentang 3-4 tahun untuk jenjang Strata 1 (S1) di Universitas Tanjungpura.

Banyak faktor yang mempengaruhi lama masa studi mahasiswa, baik faktor internal maupun faktor eksternal. Karakteristik mahasiswa yang beragam juga menjadikan lama masa studi mahasiswa menjadi beragam pula. Oleh karena itu, perguruan tinggi harus dapat melakukan evaluasi terhadap mahasiswanya. Evaluasi tersebut dilakukan guna meminimalkan faktor-faktor yang menjadi penghalang atau penghambat ketepatan waktu lulus mahasiswa. Hal tersebut menjadi sebuah alasan bagi setiap perguruan tinggi agar dapat mengambil keputusan yang tepat.

Pengambilan keputusan yang tepat oleh suatu perguruan tinggi dalam hal kelulusan mahasiswa ini guna mengevaluasi penyebab atau faktor-faktor yang menyebabkan mahasiswa lulus tidak tepat waktu. Selanjutnya perguruan tinggi bisa membuat rekomendasi untuk kedepannya berdasarkan faktor-faktor yang telah diketahui. Pengambilan keputusan yang tepat memerlukan informasi yang cukup agar bisa menganalisis lebih lanjut mengenai hal tersebut.

Informasi yang menjadi landasan dalam pengambilan keputusan bisa diperoleh dengan melakukan analisis terhadap data-data yang tersimpan di sistem informasi akademik (salah satu contoh datanya adalah data kelulusan mahasiswa). Perguruan tinggi bisa melakukan analisis terhadap data kelulusan mahasiswa sehingga dapat mengetahui sejak dini tingkat kelulusan mahasiswanya. Oleh karena itu, dibutuhkan teknik analisis data yang tepat untuk mengekstraksi informasi-informasi yang tersembunyi dari data kelulusan ini.

Menurut Teknik analisis secara manual tentunya tidak dapat digunakan untuk mengekstrak informasi dari data yang berjumlah besar. Oleh karena itu, teknik *data mining* merupakan pilihan yang tepat untuk digunakan dalam mengekstrak informasi dari tumpukan data tersebut. *Data mining* adalah cabang ilmu yang dapat membantu mempermudah dalam pengambilan keputusan dalam analisis dan mengekstraksi data (Asriningtias dan Mardhiyah, 2014). Selain itu, *data mining* juga dikenal sebagai *Knowledge Discovery in Databases* (KDD). Kegiatan yang meliputi beberapa proses mulai dari pengumpulan, pemakaian data, historis untuk menemukan keteraturan pola atau hubungan dalam set data berukuran besar, itulah yang disebut KDD (Yosepta dan Aprilianto, 2017).

Penggunaan teknik *data mining* sebagai teknik analisis diharapkan bisa memberikan informasi maupun pengetahuan yang berharga yang sebelumnya tersembunyi di dalam gudang data. Berdasarkan penelitian Mustafa, Ramadhan, dan Thenata (2019) *data mining* memiliki enam teknik analisis yaitu: deskripsi, pengelompokkan, klasifikasi, prediksi, dan estimasi. Pada penelitian ini teknik yang digunakan adalah teknik *data mining* klasifikasi. Menurut Yosepta dan Aprilianto (2017) klasifikasi merupakan proses menemukan model atau fungsi yang dapat membedakan kelas data yang memiliki tujuan memprediksi objek yang belum diketahui label kelas tersebut.

Berdasarkan uraian tersebut, maka penelitian ini akan menggunakan teknik klasifikasi dengan metode *Emerging Pattern* untuk menemukan pola kelulusan mahasiswa di Program Studi Statistika Universitas Tanjungpura. Hasil yang diharapkan dari penelitian ini yaitu membantu mempermudah manajemen dalam pengolahan data kelulusan serta dalam menganalisis dan memprediksi kelulusan mahasiswa. Kemudian hasil tersebut dapat digunakan sebagai rekomendasi pada saat pengambilan kebijakan bagi mahasiswa yang diprediksikan tidak bisa lulus tepat waktu sebagai tindakan preventif terhadap hal tersebut.

## **1.2 Rumusan Masalah**

Berdasarkan latar belakang yang diuraikan di atas, rumusan masalah yang dikembangkan adalah bagaimana penerapan teknik *data mining* klasifikasi dalam

analisis pola pada data kelulusan mahasiswa di Program Studi Statistika Universitas Tanjungpura dengan metode *Emerging Pattern*.

### **1.3 Tujuan Penelitian**

Berdasarkan rumusan masalah yang diambil, tujuan penelitian ini adalah untuk menerapkan teknik *data mining* klasifikasi dalam analisis pola pada data kelulusan mahasiswa di Program Studi Statistika Universitas Tanjungpura dengan metode *Emerging Pattern*.

### **1.4 Batasan Masalah**

Batasan masalah pada penelitian ini adalah data kelulusan mahasiswa di Program Studi Statistika Universitas Tanjungpura pada angkatan 2013-2018. Data kelulusan mahasiswa terdiri dari beberapa atribut yaitu, IPK (Indeks Prestasi Kumulatif), jenis kelamin, angkatan, dan nilai TUTEP. Alasan memilih atribut tersebut karena atribut-atribut tersebut sudah mewakili atribut-atribut lain pada data kelulusan mahasiswa.

### **1.5 Tinjauan Pustaka**

Irfan (2015) melakukan penelitian mengenai analisis pola asosiasi jalur masuk terhadap kelulusan mahasiswa dengan menggunakan metode Algoritma *Fold-Growth*; studi kasus Fakultas Sains dan Teknologi. Penelitian tersebut memiliki tujuan untuk mencari pola asosiasi antara data mahasiswa dengan data kelulusan mahasiswa apakah memiliki keterkaitan antar variabel dari kedua jenis data tersebut. Hasil dari penelitian tersebut yaitu jalur masuk melalui tes tulis lama lulusnya lebih dari 4 tahun dengan IPK berada diantara 3,00 sampai 3,49.

Farida dan Warnars (2019) melakukan penelitian prediksi pola kelulusan mahasiswa menggunakan teknik *data mining classification Emerging Pattern*; studi kasus Universitas Mercu Buana. Penelitian tersebut bertujuan untuk memprediksi pola dan menganalisa tingkat kelulusan mahasiswa. Hasil yang diperoleh dari penelitian tersebut adalah berdasarkan hasil *mining* dengan atribut jenis kelamin didapatkan persentase kelulusan tepat waktu didominasi oleh mahasiswa dengan jenis kelamin perempuan, untuk hasil *mining* dengan atribut

angkatan diperoleh bahwa mahasiswa angkatan 2009 memiliki pola kelulusan yang lebih menarik dengan persentase dan nilai *confidence* sebesar 72,5% pada kelas tidak tepat waktu. Nilai tersebut lebih besar jika dibandingkan dengan pola kelulusan mahasiswa angkatan 2010 dan 2011.

Penelitian lainnya dilakukan oleh Asriningtias dan Mardhiyah (2014). Penelitian tersebut yaitu tentang aplikasi *data mining* untuk menampilkan informasi kelulusan mahasiswa; studi kasus Program Studi Teknik Informatika Universitas Teknologi Yogyakarta. Hasil yang diperoleh dari penelitian tersebut yaitu semakin tinggi nilai *confidence* dan *support* maka semakin kuat nilai hubungan antar atribut.

Yosepta dan Aprilianto (2017) melakukan penelitian tentang analisa pola kelulusan mahasiswa pada STMIK Asia Malang dengan menggunakan Algoritma *Iterative Dichotomiser 3* (ID3). Hasil penelitian tersebut yaitu faktor yang paling berpengaruh terhadap kelulusan mahasiswa adalah indeks prestasi semester 5. Proses *testing* dengan 163 data yang diambil secara acak menghasilkan akurasi sebesar 99,39% kemudian setelah melakukan proses *testing* dengan 69 data baru menghasilkan akurasi sebesar 100%.

Nuruliyani dan Warnars (2019) melakukan penelitian tentang *prototype data mining* pola jabatan fungsional dosen menggunakan teknik *Emerging Pattern*: studi kasus Universitas Mercu Buana. Penelitian tersebut bertujuan untuk mencari pola jabatan fungsional dosen di Universitas Mercu Buana. Adapun hasil dari penelitian tersebut yaitu nilai *confidence* untuk dataset jenis kelamin pada *itemset* laki-laki dan perempuan sebesar 50% dan dataset jabatan fungsional dosen yang nilai *confidencenya* lebih besar yaitu lektor.

## **1.6 Metodologi Penelitian**

Pada penelitian ini, data yang digunakan adalah data kelulusan mahasiswa Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Tanjungpura dari tahun 2013-2018. Metodologi yang digunakan dalam penelitian ini adalah teknik *data mining* klasifikasi dengan *Emerging Pattern*. Analisis dalam penelitian ini menggunakan beberapa atribut dari data

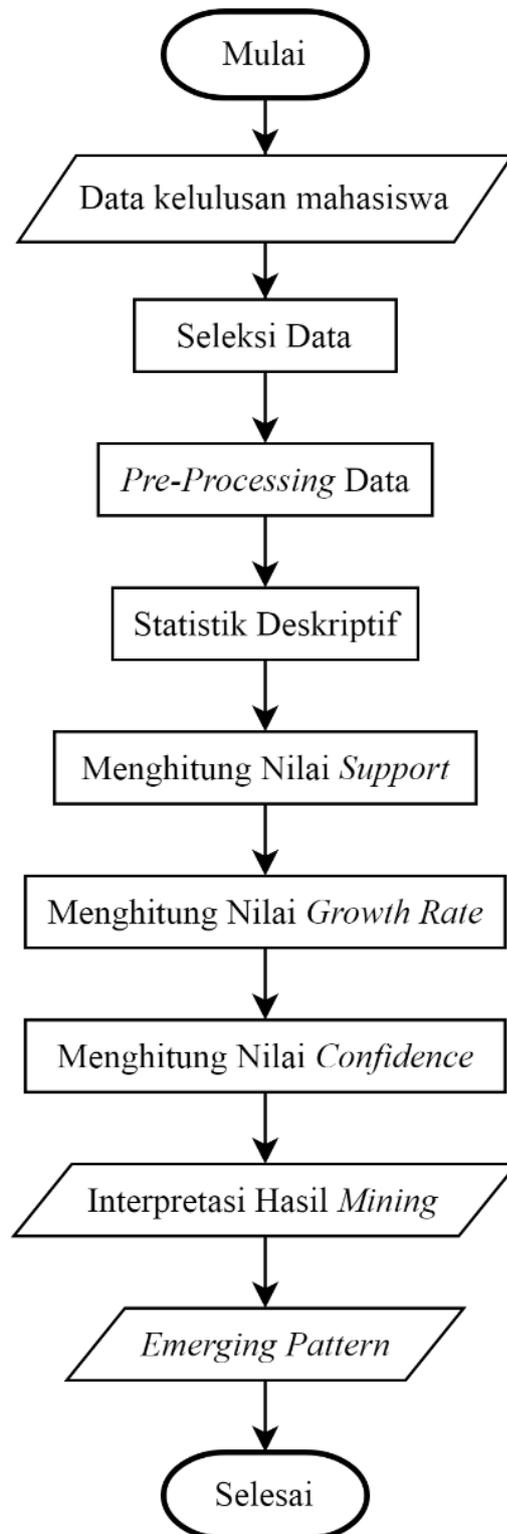
kelulusan mahasiswa. Atribut yang digunakan antara lain jenis kelamin, angkatan, nilai TUTEP, dan IPK.

Dalam memulai proses *Emerging Pattern* terlebih dahulu menginput data kelulusan mahasiswa yang sudah didapatkan. Setelah data kelulusan mahasiswa terinput selanjutnya dilakukan seleksi data. Seleksi data merupakan proses pemilihan atribut-atribut yang akan digunakan untuk dianalisis polanya.

Langkah selanjutnya, dilakukan *pre-processing* data. *Pre-processing* data yang dilakukan dalam proses ini adalah membuat atribut yang sebelumnya berupa data kuantitatif menjadi kualitatif. Sebagai contoh, pada atribut masa studi akan dibuat ke dalam bentuk kelas tepat waktu dan tidak tepat waktu. Atribut-atribut yang diubah adalah atribut nilai TUTEP dan atribut masa studi.

Langkah berikutnya adalah analisis statistik deskriptif pada seluruh atribut yang telah dipilih untuk mengetahui deskripsi dari masing-masing atribut. Berikutnya, setelah dilakukan analisis deskriptif maka akan dilakukan pencarian nilai *support*, *growth rate*, dan *confidence*. Pertama, dilakukan pencarian nilai *support* dataset dari *itemset* untuk atribut pertama. Selanjutnya mencari nilai *growth rate* dan *confidence*. Langkah-langkah tersebut dilakukan untuk seluruh atribut hingga atribut terakhir. Setelah nilai *support*, *growth rate*, dan *confidence* masing-masing atribut diperoleh maka langkah selanjutnya yaitu menginterpretasikan hasil *mining*.

Langkah terakhir dalam penelitian ini adalah *Data mining* dengan *Emerging Pattern* dimana pada proses ini merupakan proses evaluasi pola dari hasil *mining* yang diperoleh sebelumnya. Pada tahap ini dilakukan eksplorasi analisis klasifikasi *Emerging Pattern* untuk menganalisis pola dari dataset yang digunakan. Evaluasi pola dilakukan dengan cara melihat nilai *confidence* terbesar dari masing-masing *itemset*.



**Gambar 1.1** Flowchart Emerging Pattern