

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang Penelitian

Pengelolaan yang profesional menjadi tuntutan yang harus dipenuhi oleh para pengelola perguruan tinggi. Meningkatnya persaingan dalam hal penyelenggaraan proses pendidikan menjadi salah satu alasan perlunya pengelola untuk selalu memperhatikan fenomena-fenomena yang terjadi baik di lingkungan *internal* maupun *eksternalnya*. Perguruan tinggi dituntut untuk semakin efisien dalam pengelolaannya. Tuntutan ini juga dirasakan oleh pengelola program studi yang menjadi ujung tombak pelaksanaan proses belajar mengajar. Perguruan tinggi sebagai institusi pendidikan menjadi bagian penting dalam masyarakat dan memiliki peran dalam pertumbuhan dan pembangunan bangsa. Dalam bidang akademik, perguruan tinggi juga berperan dalam mengontrol, mengevaluasi serta prediksi prestasi akademik mahasiswanya.

Mengenali profil mahasiswa dengan lebih seksama dapat memberi petunjuk tentang potensi mahasiswa untuk berprestasi secara akademik. Dengan mengenali pola-pola profil mahasiswa, pihak perguruan tinggi, terutama dosen wali yang langsung berhubungan dalam pembimbingan mahasiswa dapat memanfaatkan informasi tersebut untuk melakukan antisipasi dalam langkah pembimbingan dan pengarahan mahasiswa agar dapat mengoptimalkan prestasi dan meminimalisir

potensi masalahnya (Sartikha, Winda Sari, & Jannah, 2016).

Mahasiswa sebagai salah satu aspek penting dalam evaluasi keberhasilan penyelenggaraan program studi perlu mendapatkan perhatian yang serius untuk memperoleh kepercayaan stakeholder dalam menilai dan menetapkan penggunaan lulusannya. Menurut Aziz, dkk (2018:2244) prestasi belajar mahasiswa merupakan salah satu indikator dalam menghitung kinerja mahasiswa dan efektif proses belajar mengajar, yang sekaligus dapat digunakan untuk meningkatkan citra perguruan tinggi. Salah satu indikator tinggi rendah dalam menghitung kinerja mahasiswa adalah Indeks Prestasi Kumulatif (IPK). Prestasi akademik mahasiswa dapat didasarkan pada berbagai faktor seperti kepribadian, lingkungan sosial serta psikologi dari mahasiswa tersebut.

Perguruan tinggi sebagai institusi pendidikan telah memiliki data akademik dan administrasi dalam jumlah yang sangat besar, namun hanya sebagian kecil data tersebut dimanfaatkan, khususnya dalam penyusunan evaluasi diri (Marthasari, 2017). Data akademik mahasiswa merupakan data yang dihimpun dari hasil kegiatan proses belajar mengajar selama mengikuti studi di perguruan tinggi. Dalam beberapa tahun terakhir, Universitas Pendidikan Ganesha telah menggunakan sistem informasi yang memungkinkan penghimpunan data dalam jumlah besar, baik data pribadi mahasiswa, data hasil ujian calon mahasiswa baru, data keuangan, data prestasi mahasiswa, maupun data riwayat dan hasil studi mahasiswa. Jumlah data yang banyak ini membuka peluang untuk dihasilkan informasi yang berguna bagi pihak universitas dan menjadi modal cukup penting untuk memperoleh pengetahuan-pengetahuan yang dapat menjawab pertanyaan-pertanyaan pengelola perguruan tinggi terkait permasalahan yang dihadapi.

Penggalian informasi pada sebuah data berukuran besar (mempunyai jumlah *record* dan jumlah *field* yang cukup banyak) tidak dapat dilakukan dengan mudah. Salah satu alat bantu yang dapat digunakan untuk menemukan pengetahuan/informasi yang tersembunyi dalam *database* adalah teknologi penambangan data (*data mining*). Teknik penambangan data mampu mengekstraksi pengetahuan yang tersembunyi dari data, dimana pengetahuan tersebut selanjutnya disajikan dalam format yang mudah dimengerti oleh pengguna. Metode penambangan data telah digunakan dalam analisis data pendidikan yang bidangnya dikenal dengan istilah *education data mining* (EDM). EDM menerapkan berbagai metode antara lain *decision tree*, *apriori*, *k-means*, *neural network*, dan metode lainnya. Luaran dari EDM dapat digunakan oleh para pemangku kepentingan di dunia pendidikan, termasuk pendidikan tinggi dalam pengambilan keputusan. Data yang dianalisis oleh EDM dapat berasal dari data historis dan operasional dari basis data perguruan tinggi, berupa data pribadi maupun akademik mahasiswa (Marthasari, 2017).

Dalam kasus segmentasi, teknik *clustering* dapat digunakan untuk mengevaluasi kinerja akademik melalui berbagai atribut yang mendukung prestasi akademik mahasiswa, sehingga pihak akademik universitas dapat memantau perkembangan prestasi akademik mahasiswa selama kuliah. Selain itu dapat dijadikan acuan dalam menentukan kebijakan yang tepat untuk menentukan metode pengajaran yang sesuai dengan kemampuan mahasiswa. Berdasarkan hasil observasi penelitian yang dilakukan maka teknik *clustering* yang paling tepat adalah metode *k-means*.

Metode *k-means* merupakan metode *clustering* yang paling sederhana dan umum. Hal ini dikarenakan *k-means* mempunyai kemampuan mengelompokkan

data dalam jumlah yang cukup besar dengan waktu komputasi yang relatif cepat dan efisien (Suyanto, 2017). Algoritma *k-means* digunakan karena penentuan jumlah *cluster* dapat disesuaikan dengan kebutuhan, namun *k-means* mempunyai kelemahan yang diakibatkan oleh penentuan pusat awal *cluster* dan jumlah *cluster* awal. Hasil klaster yang terbentuk dari metode *k-means* ini sangatlah tergantung pada inisialisasi nilai pusat awal klaster yang diberikan. Hal ini menyebabkan hasil klaster berupa solusi yang sifatnya *local optimal* (Widiarini & Wahono, 2015).

Untuk mengatasi permasalahan tersebut maka terdapat pengembangan algoritma *k-means* yaitu *dynamic k-means clustering*, yang memiliki proses untuk mencari jumlah *cluster* tanpa harus menduga asumsi jumlah *cluster*. *Dynamic k-means clustering* merupakan pengembangan *k-means* untuk mengecek ulang kualitas *cluster* pada setiap iterasi, memungkinkan terjadinya perubahan dalam jumlah *cluster* untuk memenuhi keabsahan kualitas *cluster*. Aplikasi algoritma klaster dinamis bertujuan untuk meningkatkan kualitas *cluster* sehingga menghasilkan angka yang optimal dari *cluster*. Penelitian pengembangan metode *k-means* telah dilakukan sebelumnya oleh (Ahamed & Hareesha, 2012) dan (Widiarini & Wahono, 2015) dengan mengusulkan algoritma klaster dinamik pada algoritma *k-means* dalam menetapkan jumlah *cluster* ( $k$ ) agar dapat menghasilkan kualitas *cluster* yang optimal, sehingga memberikan hasil pemetaan potensial lebih baik dan tepat. Namun penelitian pengembangan ini masih memiliki kekurangan terkait komputasi dalam pemrosesan data. Algoritma *dynamic k-means* memiliki kemampuan untuk mencari jumlah *cluster* ideal, namun terdapat kekurangan dalam penentuan titik *centroid* (pusat *cluster*) yang masih dipilih secara acak. Sehingga kesalahan penentuan *centroid* awal akan mempengaruhi jumlah proses iterasi dan

waktu komputasi.

Berdasarkan hal tersebut, metode klustering *k-means* dikolaborasikan dengan metode pendekatan untuk penentuan pusat awal *cluster* dan diharapkan dapat mengurangi waktu komputasi untuk *dataset* yang besar. Selain dapat menentukan jumlah klaster ideal, pemilihan algoritma klaster dinamis disebabkan pula jumlah data kinerja akademik universitas yang dihimpun semakin berkembang dan memiliki karakteristik data yang beragam. Metode yang akan diterapkan dalam penelitian ini adalah inisialisasi *centroid* menggunakan metode *mean*. Kinerja dari algoritma klustering akan dibandingkan dalam hal kemurnian, menormalkan informasi timbal balik dan waktu yang diambil untuk membentuk sebuah *cluster*. Data kinerja akademik yang dikumpulkan, akan dikelompokkan sesuai dengan karakteristik yang serupa dan membentuk klaster. Dengan menggunakan *dynamic k-means clustering*, peneliti mencoba untuk mengekstrak pengetahuan yang dapat menggambarkan kinerja akademik mahasiswa pada akhir semester selama mengikuti perkuliahan di Universitas Pendidikan Ganesha. Hasil ekstraksi pengetahuan diharapkan dapat digunakan untuk memberikan informasi persebaran kemampuan kinerja akademik mahasiswa dilihat dari data perkembangan sebelum dan setelah kuliah di Undiksha serta informasi awal persiapan promosi penerimaan mahasiswa baru. Selain itu diharapkan dapat membantu mahasiswa yang membutuhkan perhatian khusus serta memberikan informasi awal kepada instruktur dalam meningkatkan pengajaran, mengevaluasi dan meningkatkan sistem pembelajaran, meningkatkan prestasi mahasiswa, serta manfaat lainnya.

## 1.2 Identifikasi Masalah

Berdasarkan latar belakang diatas, maka dapat diidentifikasi beberapa permasalahan sebagai berikut.

1. Pengolahan data akademik mahasiswa di Undiksha masih belum optimal, terutama dalam prediksi dan pengolahan data untuk memahami perkembangan kinerja akademik setiap mahasiswa berdasarkan data-data yang tersimpan dalam Sistem Informasi Akademik.
2. Data akademik mahasiswa Undiksha belum pernah dievaluasi untuk mengetahui persebaran kemampuan mahasiswa.
3. Metode *k-means* sangat tergantung pada inialisasi nilai pusat awal *cluster* dan berpengaruh pada komputasi dalam pemrosesan data. Sementara itu pada pengembangan algoritma kluster dinamis pada *k-means* memiliki kelebihan dalam kemampuan untuk mencari jumlah kluster terbaik, namun memiliki kekurangan dalam penentuan titik *centroid* (pusat *cluster*).

## 1.3 Pembatasan Masalah

Dengan luasnya cakupan yang dapat terkait dengan penelitian ini, maka terdapat batasan-batasan yang perlu diberlakukan sebagai berikut.

1. Objek penelitian ini adalah mahasiswa program sarjana dan diploma Universitas Pendidikan Ganesha angkatan 2017.
2. Dalam penelitian yang mendasar adalah teknik *data mining* saja tidak sampai pembangunan data *warehouse*.
3. Aplikasi yang digunakan dalam penelitian ini adalah berbantuan *PHP* dan *MySQL*.

#### 1.4. Rumusan Masalah

Berdasarkan latar belakang sebelumnya, dapat dirumuskan beberapa permasalahan yang akan dibahas dalam penelitian ini sebagai berikut:

1. Bagaimana implementasi metode kluster dinamis dengan inisialisasi *centroid* awal berbasis *mean* dalam klasterisasi kinerja akademik mahasiswa?
2. Bagaimana akurasi pengembangan metode kluster dinamis pada *k-means* dengan inisialisasi *centroid* awal berbasis *mean* dalam klasterisasi kinerja akademik mahasiswa?
3. Bagaimana analisis kinerja akademik mahasiswa berdasarkan metode kluster dinamis dengan inisialisasi *centroid* awal berbasis *mean* yang terbentuk?

#### 1.5 Tujuan Penelitian

Berdasarkan rumusan masalah diatas, maka adapun tujuan yang diharapkan dapat tercapai dari penelitian adalah sebagai berikut:

1. Mengetahui implementasi metode kluster dinamis dengan inisialisasi *centroid* awal berbasis *mean* dalam klasterisasi kinerja akademik mahasiswa.
2. Mengetahui akurasi metode kluster dinamis dengan inisialisasi *centroid* awal berbasis *mean* dalam klasterisasi kinerja akademik mahasiswa.
3. Menghasilkan hasil analisis kinerja akademik mahasiswa berdasarkan metode kluster dinamis dengan inisialisasi *centroid* awal berbasis *mean* yang terbentuk.

## 1.6 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

### 1. Manfaat Teoritis

Hasil penelitian ini diharapkan dapat menjadi referensi atau masukan bagi perkembangan teknologi informasi dan komunikasi serta menambah kajian ilmu pengetahuan khususnya dalam bidang data mining. Selain itu melalui penelitian ini diharapkan dapat menghasilkan metode dalam menentukan jumlah *cluster* ideal dan penentuan pusat *cluster* awal sehingga perhitungan *k-means* lebih optimal.

### 2. Manfaat Praktis

Hasil penelitian ini diharapkan dapat menjadi masukan untuk perguruan tinggi dalam beberapa hal seperti berikut:

- a. Dapat dijadikan sebagai model untuk mengevaluasi kinerja akademik mahasiswa Universitas Pendidikan Ganesha.
- b. Dapat dijadikan masukan bagi bagian akademik universitas, fakultas, prodi, pembimbing akademik mahasiswa, serta dosen untuk mengetahui perkembangan mahasiswa selama menuntut ilmu di Universitas Pendidikan Ganesha, sehingga bisa dijadikan acuan dalam pembimbingan dan pendampingan mahasiswa dalam meningkatkan prestasi akademik mahasiswa.