

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Penetapan biaya Uang Kuliah Tunggal (UKT) telah menjadi kebijakan penting dalam sistem pendidikan tinggi di Indonesia, yang bertujuan untuk memberikan beban biaya kuliah yang sesuai dengan kemampuan finansial mahasiswa. Uang kuliah Tunggal (UKT) merupakan sistem pembayaran uang kuliah yang diterapkan oleh pemerintah pada perguruan tinggi negeri. Setiap tahunnya nominal uang pangkal mengalami kenaikan, sehingga terasa sangat memberatkan mahasiswa dalam proses pembayaran saat menjadi mahasiswa baru (Fitri *et al.*, 2020).

Meski demikian, pelaksanaan UKT di lapangan kerap menghadapi kendala ketepatan dalam pengelompokan kategori. Kesalahan ini sebagian besar terjadi karena penilaian berbasis data sosial-ekonomi yang tidak optimal atau tidak akurat, sehingga beberapa mahasiswa dengan kemampuan ekonomi berbeda ditempatkan dalam kategori yang sama. Penetapan UKT yang tidak tepat ini bisa menimbulkan beban yang tidak adil, terutama bagi mahasiswa dari latar belakang ekonomi kurang mampu. Beberapa studi menunjukkan bahwa meskipun UKT bertujuan untuk menciptakan sistem yang lebih adil, namun dalam praktiknya masih terdapat ketimpangan dalam aksesibilitas pendidikan tinggi. Hal ini dikarenakan penentuan kategori UKT yang terkadang tidak sepenuhnya mencerminkan kemampuan finansial mahasiswa, sehingga menimbulkan beban biaya yang tidak proporsional bagi beberapa mahasiswa (Fauziah & Purnamasari, 2024). Oleh karena itu, diperlukan metode yang mampu mengelompokkan data lebih akurat untuk mendukung penetapan UKT secara adil.

*Clustering* adalah teknik analisis data yang bertujuan mengelompokkan objek data ke dalam kelompok yang seragam berdasarkan karakteristik yang sama. Metode *clustering* yang digunakan dalam penelitian ini adalah *K-Means*. Metode

*K-Means Clustering* merupakan metode dalam *data mining* yang mengelompokkan data berdasarkan kemiripan karakteristik (Fajri & Purnamasari, 2022). Proses ini memungkinkan pengelompokan data yang memiliki karakteristik serupa ke dalam satu kelompok.

*Clustering* atau pengelompokan data merupakan metode yang sering digunakan untuk mengekstrak informasi yang tersembunyi dari dataset yang besar. *Clustering* bertujuan untuk mengelompokkan data berdasarkan kemiripan atau jarak antar-data sehingga dapat membentuk pola-pola yang lebih mudah diinterpretasikan. Dalam kasus Uang Kuliah Tunggal (UKT), metode *clustering* dapat digunakan untuk memahami karakteristik mahasiswa terkait biaya pendidikan, seperti latar belakang ekonomi, pekerjaan orang tua, dan faktor-faktor lain yang memengaruhi kemampuan membayar UKT. Hasil *clustering* ini dapat menjadi dasar bagi pemerintah dan institusi pendidikan untuk merumuskan kebijakan yang lebih adil dan tepat guna dalam penetapan UKT, sehingga dapat mengurangi beban biaya pendidikan bagi mahasiswa dari berbagai latar belakang ekonomi. Salah satu algoritma *clustering* yang paling populer dan banyak digunakan adalah *K-Means Clustering*. *K-Means Clustering* adalah salah satu teknik data mining tanpa pengawasan (*unsupervised*) atau semi-supervisi (*semi-supervised*). Teknik *K-Means* bekerja dengan mengelompokkan data ke dalam *cluster* sehingga data yang memiliki karakteristik serupa digabungkan menjadi satu *cluster* yang mirip, sedangkan data yang mempunyai karakteristik berbeda dikelompokkan ke dalam *cluster* lain yang sesuai (Norshahlan *et al.*, 2023).

Metode *K-Means* dikenal sebagai salah satu teknik *clustering* yang populer berkat kemampuannya dalam mengelompokkan data berukuran besar. Kelebihan metode ini adalah kesederhanaannya dan kemudahannya untuk dipahami. *K-Means* sering digunakan karena bersifat *unsupervised*, artinya tidak memerlukan label dalam proses pengelompokannya. Pemilihan *K-Means* didasari oleh keefektifannya dalam menangani data dalam jumlah besar serta kecepatan proses komputasinya. Selain itu, metode ini fleksibel, mudah diimplementasikan, dan prinsip kerjanya sederhana serta tidak memerlukan pendekatan statistik yang kompleks (Setiaji *et al.*, 2023). Namun, metode *K-Means* memiliki kelemahan dalam penentuan titik

*centroid* awal yang dilakukan secara acak, sehingga *centroid* yang diperoleh seringkali tidak optimal (Rahmadhnai *et al.*, 2020). Kelemahan dari algoritma ini terletak pada penetapan nilai K yang optimal untuk jumlah cluster yang akan dibentuk. Meskipun metode ini dapat memproses dataset berukuran besar, namun masih kurang efektif dalam mengklasifikasikan dokumen dengan akurasi yang tinggi (Maori & Evanita, 2023). Kelemahan utama *k-means* adalah bahwa jumlah k cluster harus ditentukan sebelumnya. Hal ini menimbulkan kemungkinan bahwa jumlah cluster optimal untuk dataset tertentu tidak diketahui sebelum proses *clustering*, sehingga *k-means* akan menghasilkan hasil *clustering* yang suboptimal (Aggarwal & Reddy, 2014). Jika jumlah *cluster* yang dipilih tidak tepat, hal ini dapat menyebabkan hasil *clustering* yang kurang optimal, dengan data yang terbagi ke dalam terlalu banyak atau terlalu sedikit *cluster*, sehingga mengaburkan pola yang sebenarnya.

Keterbatasan *K-Means* dalam menentukan jumlah *cluster* mendorong penggunaan metode optimasi yang dapat membantu mengatasi masalah ini. Salah satu pendekatan yang banyak diterapkan adalah penggunaan *Particle Swarm Optimization* (PSO). *Particle Swarm Optimization* adalah salah satu algoritma evolusioner yang terinspirasi dari perilaku kolektif kawanan burung dan ikan. Algoritma ini awalnya dikembangkan untuk menyelesaikan berbagai masalah optimasi dengan memanfaatkan pergerakan partikel dalam sebuah kelompok untuk mencari solusi terbaik (Setiaji *et al.*, 2023). *Particle Swarm Optimization* (PSO) adalah salah satu *algoritme swarm intelligence* yang bersifat *meta-heuristik* dan dapat digunakan untuk mengoptimalkan titik *centroid* pada metode *K-Means*, sehingga menghasilkan *cluster* yang lebih akurat. Menurut G. Komarasamy & Amitabh penerapan PSO pada *K-Means* dapat meningkatkan kinerja dibandingkan dengan algoritme *K-Means* yang murni (Rahmadhnai *et al.*, 2020).

Salah satu studi yang dilakukan oleh Dian Werdiningsih Dwi Rahmawati, Imam Cholissodin, dan Nurudin Santoso (2019) dalam artikel berjudul "Optimasi *K-Means* untuk Pengelompokan Data Kinerja Akademik Dosen menggunakan *Particle Swarm Optimization*" menilai penerapan PSO dalam menentukan jumlah *cluster* yang optimal. Hasil penelitian menunjukkan bahwa metode PSO mampu

mengurangi kesalahan pengelompokan dan menghasilkan *cluster* yang lebih homogen, dibandingkan dengan metode *K-Means* standar. Hasil pengujian dalam penelitian ini memperoleh algoritma *K-Means* yang dipadukan dengan algoritma *Particle Swarm Optimization* lebih baik 3.28% dibandingkan dengan *algoritma K-Means* yang berdiri sendiri, dimana kualitas klaster ditentukan menggunakan metode *Silhouette Coefficient* (Rahmawati *et al.*, 2019).

Penelitian lainnya yang dilakukan oleh Indrawan Setiaji, Affandy, Ahmad Zainul Fanani (2023) berjudul "Optimasi *K-Means Clustering* Dengan Menggunakan *Particle Swarm Optimization* Untuk Menentukan Jumlah *Cluster* Pada Kanker Serviks". Penelitian ini menggunakan metode *K-Means* yang dioptimalkan dengan *Particle Swarm Optimization* (PSO) untuk menentukan jumlah *cluster* secara lebih efektif. Metode *K-Means* dianggap tidak bisa memberikan solusi optimal penentuan jumlah *clustering*, sehingga perlu untuk ditingkatkan agar mendapatkan solusi optimal. Evaluasi hasil *clustering* dilakukan menggunakan *Silhouette Coefficient* (SI) dan *Davies Bouldin Index* (DBI). Hasil penelitian menunjukkan bahwa metode PSO - *K-Means* dapat terbukti memberikan kontribusi yang signifikan dengan langsung mendapatkan hasil *clustering* yang optimum tanpa harus melakukan percobaan berulang kali dengan nilai *Silhouette Coefficient* mencapai 0,83 dan nilai *Davies Bouldien Index* mencapai 1,91 (Setiaji *et al.*, 2023).

Berdasarkan hasil penelitian sebelumnya menunjukkan potensi besar dari PSO untuk mengatasi kelemahan yang dimiliki oleh *K-Means*, khususnya dalam hal penentuan jumlah *cluster* yang optimal. Dengan menggunakan PSO untuk mengoptimalkan *K-Means Clustering*, penelitian ini diharapkan dapat memberikan solusi yang lebih baik dalam pengelompokan data terkait Uang Kuliah Tunggal (UKT). Penggunaan metode ini dapat membantu institusi pendidikan untuk mendapatkan gambaran yang lebih jelas tentang kondisi ekonomi mahasiswa, serta memberikan wawasan bagi pemerintah dan lembaga terkait untuk merumuskan kebijakan UKT yang lebih adil dan sesuai dengan kemampuan finansial mahasiswa.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana menentukan jumlah *cluster* optimal pada pengelompokan data Uang Kuliah Tunggal?
2. Bagaimana penerapan *Particle Swarm Optimization* dalam optimasi jumlah *cluster* pada metode *K-Means Clustering*?
3. Bagaimana membangun sistem yang menerapkan *Particle Swarm Optimization* untuk mengoptimalkan jumlah *cluster* pada metode *K-Means Clustering*?

## 1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah ditetapkan, tujuan penelitian ini adalah sebagai berikut:

1. Menentukan jumlah *cluster* optimal untuk data UKT melalui optimasi *K-Means* menggunakan PSO.
2. Untuk mengimplementasikan metode *Particle Swarm Optimization* dalam meningkatkan performa *K-Means clustering* dalam mengelompokkan data UKT mahasiswa.
3. Untuk membangun sistem yang menerapkan *Particle Swarm Optimization* dalam mengoptimalkan jumlah *cluster* pada metode *K-Means Clustering*.

## 1.4 Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut:

1. Penelitian ini diharapkan dapat memperkaya literatur mengenai optimasi metode *clustering*, khususnya dalam penggunaan *Particle Swarm Optimization* (PSO) pada *K-Means Clustering*. Dengan hasil yang diperoleh, penelitian ini dapat menjadi referensi bagi penelitian-penelitian selanjutnya yang terkait dengan optimasi jumlah *cluster*.

2. Hasil penelitian ini dapat diaplikasikan dalam menentukan jumlah *cluster* yang optimal pada pengelompokan UKT Mahasiswa atau data serupa, sehingga mempermudah pengambilan keputusan dalam pemetaan data yang kompleks.
3. Penelitian ini memberikan kontribusi dalam pengembangan sistem berbasis teknologi yang mampu melakukan pengelompokan data secara optimal.

### **1.5 Ruang Lingkup Dan Batasan Penelitian**

Ruang lingkup dan batasan masalah yang ditetapkan dalam penelitian ini adalah sebagai berikut:

1. Penelitian ini hanya berfokus pada penerapan metode optimasi *K-Means clustering* menggunakan *Particle Swarm Optimization* (PSO) untuk menentukan jumlah *cluster* yang optimal pada data UKT.
2. Penelitian ini membahas penggunaan PSO dalam mengoptimalkan *K-Means clustering* dan menggunakan metode elbow untuk perbandingan hasil optimasi
3. Data yang digunakan dalam penelitian ini adalah data historis UKT mahasiswa baru tahun 2024 dari Fakultas Teknik yang diambil dari UPT PUSKOM Universitas Malikussaleh untuk memastikan relevansi dan keterkinian analisis.
4. Pengujian dan evaluasi hasil *clustering* akan dilakukan menggunakan metrik evaluasi standar *Silhouette Coefficient* dan *Davies-Bouldin Indeks* untuk menilai kualitas *cluster* yang dihasilkan.