

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pendidikan merupakan fondasi utama dalam pembangunan kualitas sumber daya manusia dan kemajuan bangsa. Akses terhadap pendidikan tinggi yang merata menjadi tantangan penting, terutama bagi masyarakat dengan keterbatasan ekonomi. Di berbagai daerah, masih banyak generasi muda yang tidak melanjutkan pendidikan ke perguruan tinggi karena faktor biaya (Ihsanuddin et al., 2024). Kondisi tersebut menunjukkan bahwa aspek finansial menjadi hambatan signifikan dalam pemerataan pendidikan. Oleh karena itu, diperlukan kebijakan dan program yang mampu menjembatani kesenjangan akses tersebut. Salah satu bentuk intervensi yang umum diterapkan adalah pemberian beasiswa sebagai bantuan pendidikan. Program ini bertujuan membantu mahasiswa agar tetap dapat menyelesaikan studi tanpa terbebani kendala ekonomi.

Beasiswa merupakan bentuk bantuan finansial yang diberikan kepada mahasiswa berdasarkan kriteria tertentu, baik akademik maupun ekonomi. Selain meningkatkan motivasi belajar, beasiswa juga berfungsi sebagai instrumen pemerataan kesempatan pendidikan (Salendah et al., 2022). Dalam praktiknya, proses seleksi beasiswa sering menghadapi permasalahan ketidaktepatan sasaran. Hal ini terjadi karena proses penilaian masih dilakukan secara manual dan belum memanfaatkan analisis data secara optimal. Ketidaktepatan tersebut dapat menimbulkan ketidakadilan distribusi bantuan. Oleh sebab itu, sistem seleksi yang objektif dan berbasis data menjadi kebutuhan penting dalam pengelolaan beasiswa.

Dalam sistem pendidikan tinggi di Indonesia, Uang Kuliah Tunggal (UKT) diterapkan sebagai skema pembiayaan yang disesuaikan dengan kemampuan ekonomi mahasiswa. UKT dibagi ke dalam beberapa kelompok berdasarkan kondisi sosial ekonomi orang tua. Namun, penentuan kategori UKT sering kali menghadapi kendala dalam pengolahan data yang kompleks (Kusnaidi et al., 2022). Tanpa metode analisis

yang tepat, pengelompokan UKT dapat menghasilkan keputusan yang kurang akurat. Oleh karena itu, diperlukan pendekatan analitis untuk mengelompokkan mahasiswa secara sistematis berdasarkan indikator yang relevan.

Permasalahan distribusi bantuan UKT juga ditemukan pada beberapa perguruan tinggi yang mengalami keterlambatan serta kurang tepat sasaran dalam menentukan penerima bantuan (Kusnaidi et al., 2022). Hal ini menunjukkan bahwa proses manual memiliki keterbatasan dalam mengolah data mahasiswa yang berjumlah besar. Dengan meningkatnya volume data akademik dan ekonomi mahasiswa, pengelolaan informasi memerlukan pendekatan berbasis teknologi. Pemanfaatan teknik komputasi menjadi solusi untuk meningkatkan akurasi dan efisiensi proses seleksi. Salah satu pendekatan yang dapat diterapkan adalah data mining.

Data mining merupakan proses mengekstraksi pola dan informasi bermakna dari kumpulan data besar menggunakan teknik statistik dan machine learning. Sudarsono & Lestari (2021), menjelaskan bahwa data mining merupakan bagian dari proses Knowledge Discovery in Databases (KDD) yang meliputi seleksi, transformasi, analisis, dan evaluasi data. Pendekatan ini memungkinkan identifikasi pola tersembunyi yang sulit ditemukan melalui analisis manual. Dalam konteks pendidikan, data mining dapat digunakan untuk mendukung sistem pengambilan keputusan. Dengan demikian, penerapan data mining relevan dalam pengelompokan data penerima beasiswa.

Peran strategis data mining tidak hanya terbatas pada sektor bisnis, tetapi juga dalam transformasi sistem pengambilan keputusan berbasis data (Setiyani et al., 2025). Pemanfaatan teknik seperti klasifikasi dan clustering membantu organisasi memperoleh wawasan yang lebih akurat. Dalam bidang pendidikan, teknik clustering dapat digunakan untuk mengelompokkan mahasiswa berdasarkan karakteristik tertentu. Hal ini mendukung pengambilan keputusan yang lebih transparan dan terukur. Oleh karena itu, integrasi data mining dalam sistem seleksi beasiswa menjadi langkah inovatif yang relevan.

Clustering merupakan teknik unsupervised learning yang bertujuan mengelompokkan data berdasarkan tingkat kemiripan karakteristiknya (Yuliana et al., 2022). Teknik ini tidak memerlukan label kelas sebelumnya sehingga sesuai untuk eksplorasi pola data. Dalam pengelompokan mahasiswa, clustering dapat membantu mengidentifikasi kelompok yang memiliki kondisi serupa. Hasil pengelompokan dapat digunakan sebagai dasar rekomendasi kebijakan. Oleh sebab itu, clustering menjadi metode yang tepat dalam analisis data beasiswa berbasis UKT dan semester.

Algoritma K-Means merupakan salah satu metode clustering yang paling banyak digunakan. Harahap & Rambe (2021), memperkenalkan algoritma ini sebagai metode partisi yang membagi data ke dalam sejumlah cluster berdasarkan centroid. K-Means bekerja dengan menghitung jarak antar data menggunakan pendekatan Euclidean distance. Proses iteratif dilakukan hingga centroid stabil dan tidak berubah lagi. Metode ini dikenal efisien serta mudah diimplementasikan. Oleh karena itu, K-Means banyak diterapkan dalam berbagai penelitian pendidikan (Prima et al., 2022).

Meskipun K-Means memiliki keunggulan dalam efisiensi, algoritma ini sensitif terhadap pemilihan centroid awal. Pemilihan jumlah cluster yang tidak tepat dapat memengaruhi kualitas hasil pengelompokan (Orisa, 2022). Oleh karena itu, evaluasi cluster menjadi tahapan penting dalam proses analisis. Metode evaluasi internal seperti Elbow, Davies-Bouldin Index, dan Silhouette Index digunakan untuk menentukan jumlah cluster optimal. Evaluasi ini membantu menghasilkan cluster yang lebih representatif terhadap pola data.

Metode Elbow mengukur variasi dalam cluster melalui nilai Sum of Square Error (SSE) untuk menentukan titik optimal jumlah cluster (Orisa, 2022). Sementara itu, Davies-Bouldin Index mengukur rasio antara kohesi dan separasi cluster. Silhouette Index menilai kedekatan objek terhadap cluster sendiri dibandingkan cluster lain. Penggunaan kombinasi metode evaluasi ini meningkatkan validitas hasil clustering. Oleh sebab itu, evaluasi cluster perlu diterapkan dalam penelitian berbasis K-Means.

Sebelum proses clustering dilakukan, data harus melalui tahap preprocessing dan normalisasi. Perbedaan rentang nilai antar atribut dapat menyebabkan bias dalam perhitungan jarak (Kusnaldi et al., 2022). Normalisasi bertujuan menyamakan skala nilai sehingga setiap atribut memiliki kontribusi yang seimbang. Metode normalisasi seperti Min-Max Scaling dan Decimal Scaling sering digunakan dalam penelitian clustering. Tahap ini sangat penting untuk menghasilkan model yang lebih akurat.

Penelitian terdahulu menunjukkan bahwa penerapan normalisasi sebelum K-Means meningkatkan kualitas hasil pengelompokan (Kusnaldi et al., 2022). Tanpa normalisasi, atribut dengan rentang nilai besar cenderung mendominasi perhitungan jarak. Oleh karena itu, transformasi data menjadi langkah esensial dalam proses analisis. Dalam konteks penelitian ini, variabel semester dan UKT memiliki rentang nilai berbeda sehingga memerlukan normalisasi terlebih dahulu.

Beberapa penelitian sebelumnya telah menerapkan K-Means dalam seleksi beasiswa dengan berbagai variabel. Namun, sebagian besar menggunakan banyak atribut kompleks seperti IPK, penghasilan orang tua, dan prestasi (Kusnaldi et al., 2022). Penelitian tersebut belum secara spesifik menekankan kombinasi sederhana antara semester dan UKT sebagai variabel utama. Hal ini membuka peluang untuk melakukan penelitian dengan fokus variabel yang lebih terarah dan relevan.

Variabel semester dalam penelitian ini merepresentasikan tingkat keberlanjutan studi mahasiswa. Semester menunjukkan posisi akademik mahasiswa dalam siklus pendidikan. Mahasiswa pada semester akhir memiliki kebutuhan pembiayaan yang berbeda dibandingkan mahasiswa awal. Oleh karena itu, semester menjadi indikator penting dalam pengambilan keputusan bantuan pendidikan (Salendah et al., 2022). Sementara itu, UKT mencerminkan tingkat beban biaya pendidikan yang ditanggung mahasiswa. Semakin tinggi kategori UKT, semakin besar biaya yang harus dibayarkan. Penggunaan UKT sebagai variabel penelitian memberikan gambaran kondisi ekonomi mahasiswa secara langsung (Kusnaldi et al., 2022). Oleh karena itu, kombinasi semester dan UKT menjadi pendekatan yang relevan dalam pengelompokan penerima beasiswa.

Berdasarkan kajian teori dan penelitian terdahulu, terlihat bahwa metode clustering telah banyak diterapkan dalam bidang pendidikan. Namun, masih terdapat kesenjangan dalam penerapan metode tersebut pada konteks daerah tertentu seperti Kabupaten Konawe Selatan. Sebagian besar penelitian dilakukan pada tingkat universitas tertentu dan belum mengarah pada pengelompokan berbasis wilayah. Hal ini menunjukkan adanya research gap yang perlu diteliti lebih lanjut.

Selain itu, penelitian sebelumnya belum banyak mengintegrasikan evaluasi cluster secara komprehensif dalam analisis beasiswa. Padahal, evaluasi cluster sangat penting untuk memastikan kualitas hasil pengelompokan (Orisa, 2022). Oleh karena itu, penelitian ini mengombinasikan K-Means dengan metode evaluasi cluster untuk memperoleh hasil optimal.

Penggunaan pendekatan berbasis data diharapkan mampu meningkatkan transparansi dalam distribusi beasiswa. Sistem yang objektif akan mengurangi potensi subjektivitas dalam proses seleksi. Selain itu, pemanfaatan data mining mendukung kebijakan pendidikan berbasis bukti (Setiyani et al., 2025). Pendekatan ini sejalan dengan transformasi digital dalam pengelolaan pendidikan.

Dalam konteks Kabupaten Konawe Selatan, penerapan metode K-Means Clustering diharapkan dapat menjadi salah satu pendekatan analitis dalam memahami karakteristik mahasiswa penerima program beasiswa daerah. Program beasiswa tersebut merupakan program prioritas pemerintah yang diperuntukkan bagi seluruh mahasiswa yang berasal dari Kabupaten Konawe Selatan tanpa membedakan kondisi ekonomi. Oleh karena itu, penelitian ini tidak dimaksudkan untuk menentukan kelayakan penerima beasiswa, melainkan untuk melakukan pengelompokan data berdasarkan kesamaan karakteristik Semester (SMT) dan Uang Kuliah Tunggal (UKT).

Variabel Semester digunakan untuk menggambarkan tahapan studi mahasiswa, sedangkan variabel UKT merepresentasikan variasi beban biaya pendidikan yang ditanggung mahasiswa. Penggunaan kedua variabel ini memungkinkan dilakukan analisis distribusi mahasiswa secara lebih sistematis dan terstruktur, sehingga dapat memberikan

gambaran mengenai pola tahapan studi dan struktur biaya pendidikan mahasiswa penerima beasiswa di Kabupaten Konawe Selatan.

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk menerapkan metode K-Means Clustering dalam pengelompokan data penerima beasiswa di Kabupaten Konawe Selatan dengan fokus pada variabel Semester dan UKT. Proses analisis dilakukan melalui tahapan praproses data, normalisasi, penentuan jumlah kluster optimal, serta evaluasi kualitas kluster untuk memastikan hasil yang diperoleh memiliki tingkat kekompakan dan pemisahan yang baik. Penelitian ini diharapkan dapat memberikan kontribusi akademik dalam pengembangan penerapan data mining pada bidang pendidikan serta kontribusi praktis bagi pemerintah daerah dalam mendukung evaluasi dan perencanaan program beasiswa secara lebih terstruktur, objektif, dan berbasis data.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang yang telah dijelaskan sebelumnya, maka permasalahan yang akan dikaji dalam penelitian ini dapat dirumuskan sebagai berikut:

1. Bagaimana proses pengelompokan data penerima beasiswa di Kabupaten Konawe Selatan menggunakan metode K-Means berdasarkan variabel Semester dan UKT?
2. Berapa jumlah cluster terbaik yang dihasilkan serta bagaimana kualitas hasil pengelompokan tersebut berdasarkan evaluasi SSE dan Silhouette Score?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah ditetapkan, penelitian ini bertujuan untuk:

1. Mengimplementasikan metode K-Means Clustering dalam proses pengelompokan data penerima beasiswa di Kabupaten Konawe Selatan dengan memanfaatkan variabel Semester (SMT) dan Uang Kuliah Tunggal (UKT) sebagai dasar analisis.

2. Menentukan jumlah cluster yang paling sesuai menggunakan pendekatan Elbow Method, serta menilai kualitas hasil pengelompokan berdasarkan pengukuran Sum of Squared Error (SSE) dan Silhouette Score.

1.4 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah:

1.4.1 Manfaat Teoritis

Penelitian ini diharapkan dapat memberikan tambahan pengetahuan dan pemahaman dalam bidang data mining, terutama dalam penerapan metode clustering dengan algoritma K-Means untuk mengelompokkan data mahasiswa penerima beasiswa.

1.4.2 Manfaat Praktis

1. Bagi Pemerintah Kabupaten Konawe Selatan, penelitian ini dapat membantu dalam melakukan analisis dan segmentasi data mahasiswa penerima beasiswa berdasarkan karakteristik Semester (SMT) dan besaran UKT. Hasil pengelompokan ini dapat digunakan sebagai bahan evaluasi program beasiswa daerah, sehingga pemerintah dapat memahami distribusi mahasiswa berdasarkan tahapan studi dan variasi beban biaya pendidikan secara lebih terstruktur dan berbasis data.
2. Bagi instansi pengelola program beasiswa, hasil penelitian ini dapat menjadi dasar dalam melakukan perencanaan dan pengelolaan anggaran pendidikan secara lebih sistematis, serta membantu dalam memetakan kebutuhan mahasiswa berdasarkan kelompok karakteristik yang terbentuk.
3. Bagi institusi pendidikan dan peneliti selanjutnya, penelitian ini dapat menjadi referensi dalam penerapan metode data mining, khususnya algoritma K-Means Clustering, untuk analisis data pendidikan serta pengembangan variabel atau metode yang lebih komprehensif di masa mendatang.

1.5 Ruang Lingkup Masalah

Agar penelitian lebih terarah dan sesuai dengan tujuan yang telah ditetapkan, maka ruang lingkup masalah dalam penelitian ini dibatasi sebagai berikut:

1. Penelitian ini berfokus pada pengelompokan data penerima beasiswa di Kabupaten Konawe Selatan menggunakan metode K-Means Clustering.
2. Data yang digunakan merupakan data mahasiswa penerima beasiswa yang tersedia pada instansi terkait di Kabupaten Konawe Selatan.
3. Variabel yang digunakan dalam proses analisis hanya meliputi Semester (SMT) dan Uang Kuliah Tunggal (UKT).
4. Penelitian ini hanya melakukan proses pengelompokan data (clustering) dan tidak membahas penentuan kelayakan atau keputusan akhir penerima beasiswa.

1.6 Kebaharuan

Kebaharuan dalam penelitian ini terletak pada pendekatan yang digunakan dalam pengelompokan data penerima beasiswa, yaitu:

1. Penelitian ini mengombinasikan metode K-Means Clustering dengan metode Elbow untuk menentukan jumlah cluster yang optimal.
2. Evaluasi kualitas hasil pengelompokan dilakukan menggunakan Silhouette Score untuk memastikan tingkat pemisahan antar cluster.
3. Variabel penelitian difokuskan pada Semester (SMT) dan Uang Kuliah Tunggal (UKT), yang mencerminkan kondisi akademik dan beban biaya pendidikan mahasiswa.
4. Penelitian dilakukan pada konteks wilayah Kabupaten Konawe Selatan, sehingga memberikan kontribusi analisis beasiswa pada tingkat daerah.