

Penerapan Metode K-Means pada Data Kinerja Mahasiswa untuk Menentukan Kelompok Prestasi Akademik (Studi Kasus : Universitas ABC)

Allesandro Umbu Balla Rundi^{1)*}, Erni Widarti²⁾

¹⁾²⁾Sistem Informasi Kota Cerdas, Fakultas Teknik, Universitas Tunas Pembangunan Surakarta
E-mail: ^{1)*} f0220001_allesandroumbuballarundi@student.utp.ac.id, ²⁾ erni.widarti@lecture.utp.ac.id

*Corresponding

Abstrak

Penelitian ini bertujuan untuk menerapkan metode K-Means pada data kinerja mahasiswa untuk menentukan kelompok prestasi akademik di Universitas ABC. Metode K-Means digunakan untuk mengelompokkan mahasiswa berdasarkan kinerja akademik mereka. Data yang digunakan dalam penelitian ini mencakup 30 mahasiswa yang diidentifikasi oleh nama mereka. Hasil penelitian menunjukkan adanya tiga kelompok prestasi akademik yang dapat diidentifikasi menggunakan metode K-Means. Cluster 1 (C1) terdiri dari 17 mahasiswa yang dapat dikategorikan sebagai "lolos" dalam kelompok prestasi akademik. Cluster 2 (C2) terdiri dari 7 mahasiswa yang dapat dikategorikan sebagai "uji coba" dalam kelompok prestasi akademik. Cluster 3 (C3) terdiri dari 6 mahasiswa yang dapat dikategorikan sebagai "tidak lolos" dalam kelompok prestasi akademik. Anggota-anggota mahasiswa dalam setiap cluster ditentukan berdasarkan analisis kinerja akademik mereka. Mahasiswa-mahasiswa dalam Cluster 1 menunjukkan kinerja yang baik dan memiliki potensi untuk melanjutkan ke program studi lanjut. Mahasiswa-mahasiswa dalam Cluster 2 menunjukkan kinerja yang sedang dan perlu uji coba lebih lanjut untuk menentukan kemampuan mereka dalam prestasi akademik. Mahasiswa-mahasiswa dalam Cluster 3 menunjukkan kinerja yang rendah dan tidak memenuhi kriteria untuk lolos dalam kelompok prestasi akademik. Penelitian ini memberikan kontribusi penting dalam mengevaluasi kinerja akademik mahasiswa di Universitas ABC dan membantu dalam menentukan kelompok prestasi akademik mereka. Hasil penelitian ini dapat digunakan sebagai pedoman dalam pengambilan keputusan terkait program studi lanjut dan pengembangan potensi akademik mahasiswa.

Kata Kunci— *K-Means, Data Kinerja Mahasiswa, Pengelompokan Evaluasi*

Abstract

This study aims to apply the K-Means method to student performance data to determine academic achievement groups at ABC University. The K-Means method is used to cluster students based on their academic performance. The data used in this study includes 30 students identified by their names. The results of the study identified three academic achievement groups using the K-Means method. Cluster 1 (C1) consists of 17 students categorized as "passed" in the academic achievement group. Cluster 2 (C2) consists of 7 students categorized as "on trial" in the academic achievement group. Cluster 3 (C3) consists of 6 students categorized as "failed" in the academic achievement group. The students in each cluster were determined based on their academic performance analysis. Students in Cluster 1 demonstrated good performance and have the potential to proceed to advanced study programs. Students in Cluster 2 showed moderate performance and require further assessment to determine their capability in academic achievement. Students in Cluster 3 exhibited low performance and did not meet the criteria to pass within the academic achievement group. This study provides valuable insights into evaluating student academic performance at ABC University and assists in determining their academic achievement groups. The results of this study can be used as a guideline in decision-making related to advanced study programs and the development of students' academic potential.

Keywords— *K-Means, Student Performance Data, Evaluation Clustering*

1. Pendahuluan

Data mining adalah salah satu metode dalam analisis data yang bertujuan untuk menemukan pola atau informasi yang berguna dari data besar dan kompleks [1]. Metode ini telah banyak digunakan dalam berbagai bidang, termasuk bisnis, ilmu sosial, kesehatan, dan pendidikan. Dalam dunia bisnis, data mining dapat membantu dalam pengambilan keputusan bisnis, memprediksi tren pasar, dan meningkatkan efisiensi operasional perusahaan. Di bidang ilmu sosial, data mining dapat membantu dalam pemodelan perilaku konsumen dan prediksi hasil pemilihan [2]. Dalam bidang kesehatan, data mining dapat membantu dalam identifikasi risiko penyakit dan pengembangan rencana perawatan yang lebih efektif. Namun, pengolahan data yang besar dan kompleks menjadi kendala dalam analisis data tersebut [3]. Oleh karena itu, dibutuhkan metode analisis data yang efektif dan efisien dalam menggali pola dan informasi dari Analisis data mining adalah teknik untuk menganalisis data dan mengidentifikasi pola atau hubungan di antara variabel-variabel yang terkait. Metode analisis data mining dapat membantu menggali informasi yang tersembunyi dalam data, sehingga dapat digunakan untuk mengambil keputusan yang lebih baik.

Dalam analisis data mining, terdapat beberapa teknik seperti klasifikasi, regresi, klustering, asosiasi, dan lain sebagainya. Salah satu teknik yang paling umum digunakan adalah klustering. Klustering adalah teknik untuk mengelompokkan objek-objek dalam sebuah data berdasarkan kesamaan atau kemiripan antara objek-objek tersebut [4]. Salah satu metode klustering yang paling umum digunakan adalah metode K-Means. Metode K-Means adalah teknik untuk mengelompokkan data dengan cara membagi data ke dalam beberapa kelompok atau kluster, di mana setiap data atau objek dalam kelompok tersebut memiliki kesamaan atau kemiripan tertentu. Metode ini dilakukan dengan cara menghitung jarak antara setiap data dengan pusat kluster atau centroid, dan membagi data ke dalam kelompok yang memiliki pusat kluster terdekat. data besar dan kompleks [5]. Penerapan metode K-Means dalam analisis data mining dapat digunakan untuk memecahkan berbagai masalah di berbagai bidang, seperti bisnis [6][7], kesehatan [8][9], pariwisata [10][11], dan Pendidikan [12][13]. Dalam pendidikan, metode K-Means dapat digunakan untuk mengelompokkan mahasiswa berdasarkan kinerja akademiknya, sehingga dapat membantu universitas dalam memberikan layanan terbaik bagi mahasiswa dan meningkatkan kualitas pendidikan di universitas.

Universitas adalah lembaga pendidikan tinggi yang memberikan pendidikan akademik dan profesional kepada mahasiswanya dalam berbagai bidang ilmu pengetahuan, teknologi, dan seni. Universitas biasanya menawarkan program sarjana, magister, dan doktor, serta berbagai program pendidikan profesional dan pengembangan keterampilan. Universitas juga memiliki peran penting dalam penelitian dan pengembangan ilmu pengetahuan dan teknologi. Banyak universitas memiliki fakultas atau departemen penelitian yang melakukan penelitian dalam berbagai bidang, seperti sains, teknologi, kedokteran, dan lain sebagainya [14]. Salah satu universitas yang memiliki visi dan misi untuk menghasilkan lulusan yang berkualitas adalah Universitas ABC.

Universitas ABC adalah sebuah universitas yang memiliki reputasi yang baik dalam bidang pendidikan dan penelitian. Terletak di sumba timur, Universitas ABC menawarkan berbagai program pendidikan tinggi, seperti program sarjana, program magister, dan program doktoral. Universitas ABC memiliki banyak mahasiswa dengan berbagai latar belakang dan kemampuan, dan fokus pada mempersiapkan mahasiswa untuk sukses di bidang akademik dan profesional. Universitas ABC menempatkan nilai-nilai seperti integritas, inovasi, kerjasama, dan keunggulan akademik sebagai inti dari pendidikan yang diberikan. Selain menawarkan program studi yang beragam, Universitas ABC juga memiliki fasilitas yang lengkap untuk mendukung kegiatan akademik dan kehidupan kampus mahasiswa. Beberapa fasilitas yang tersedia di Universitas ABC meliputi perpustakaan, laboratorium, pusat olahraga, dan asrama mahasiswa. Dalam upaya untuk meningkatkan penelitian dan pengembangan ilmu pengetahuan, Universitas ABC memiliki beberapa pusat penelitian yang fokus pada bidang-bidang tertentu, seperti teknologi informasi, kesehatan, dan lingkungan. Secara keseluruhan, Universitas ABC adalah sebuah institusi pendidikan tinggi yang berfokus pada memberikan pendidikan akademik yang berkualitas dan mempersiapkan mahasiswanya untuk sukses di dunia akademik dan profesional.

Universitas ABC memiliki permasalahan dalam menentukan kelompok prestasi akademik mahasiswanya karena memiliki banyak mahasiswa dengan berbagai latar belakang dan kemampuan. Dalam menentukan prestasi akademik, Universitas ABC harus mempertimbangkan banyak faktor, seperti

rata-rata nilai, jumlah sks yang diambil, dan faktor lain yang relevan. Klasifikasi prestasi akademik yang dilakukan secara manual dapat memakan waktu dan sumber daya manusia yang cukup banyak, dan masih terdapat kesalahan dalam pengelompokan. Oleh karena itu, diperlukan metode yang lebih efektif dan efisien dalam mengelompokkan data kinerja akademik mahasiswa.

Berdasarkan permasalahan di atas hal ini dapat menyebabkan jika pengelompokan kelompok prestasi akademik mahasiswa di Universitas ABC akan dilakukan secara manual oleh tenaga kependidikan, seperti dosen atau koordinator akademik. Proses pengelompokan ini memakan waktu dan sumber daya manusia yang cukup banyak, terutama jika Universitas ABC memiliki banyak mahasiswa. Selain itu, pengelompokan yang dilakukan secara manual dapat menghasilkan hasil yang tidak akurat karena didasarkan pada persepsi dan penilaian subjektif dari tenaga kependidikan yang berbeda-beda. Oleh karena itu tujuan dari penelitian ini adalah menerapkan metode K-means dalam mengelompokkan mahasiswa berdasarkan kinerja akademik mereka, serta menentukan kelompok prestasi akademik yang ada di Universitas ABC. Selain itu, tujuan dari penelitian ini juga untuk memberikan kontribusi dalam pengambilan keputusan akademik yang lebih tepat sasaran, pengembangan strategi pembelajaran, dan bimbingan mahasiswa di Universitas ABC, dengan mengidentifikasi perbedaan prestasi akademik di antara kelompok prestasi akademik, penelitian ini akan memberikan wawasan berharga bagi pihak akademik dalam memahami kebutuhan dan karakteristik mahasiswa pada program studi lanjutan di Universitas ABC. Hasil penelitian ini diharapkan dapat mendukung pengembangan program pendidikan yang lebih efektif dan mendukung upaya meningkatkan prestasi akademik mahasiswa.

2. Tinjauan Pustaka

Tinjauan pustaka dalam penelitian ini akan membahas beberapa penelitian terkait yang relevan dengan topik yang dibahas dalam penelitian. Penelitian yang dilakukan oleh Hendro Priyatman, Fahmi Sajid, dan Dannis Haldivany pada tahun 2019 tentang Klasterisasi Menggunakan Algoritma K-Means Clustering untuk Memprediksi Waktu Kelulusan Mahasiswa. Tujuan dari penelitian tersebut adalah mengklasterisasi mahasiswa dengan algoritma KMeans, mengklaster data-data mahasiswa berdasarkan parameter yang ada yaitu IPK dan kehadiran, dan mengetahui kemampuan algoritma K-Means Clustering dalam memprediksi waktu kelulusan mahasiswa serta memberikan sarana untuk mengetahui perkiraan waktu lulus. Metode yang digunakan dari penelitian tersebut merupakan clustering menggunakan algoritman K-Means. Hasil dari penelitian tersebut terbagi menjadi empat cluster yaitu C1 memiliki IPK 3.190 dengan kehadiran 5, C2 memiliki IPK 3.11 dengan kehadiran 4, C3 memiliki IPK 3.09 dengan kehadiran 3, dan C4 memiliki IPK 3.05 dengan kehadiran 2. Penelitian tersebut diharapkan dapat membantu pihak kampus maupun mahasiswa untuk memprediksi tingkat kelulusan tepat waktu dan untuk meningkatkan reputasi bagi pihak kampus itu sendiri dan kelulusan tepat waktu bagi mahasiswa supaya kelulusan mereka tidak terlambat, selain itu pihak kampus bisa melakukan hal-hal yang perlu dilakukan apabila mereka di prediksi lulus tidak tepat waktu seperti dengan melakukan bimbingan dan hal lainnya [15].

Penelitian yang di lakukan oleh Abu Salam, Diyan Adiatma dan , Junta Zeniarja pada tahun 2020 tentang Implementasi Algoritma K-Means dalam Pengklasteran untuk Rekomendasi Penerima Beasiswa PPA di UDINUS. Tujuan dari penelitian tersebut adalah untuk merekomendasikan penerima beasiswa dengan menggunakan algoritma k-means, hasil rekomendasi berupa penempatan data pendaftar beasiswa ke masing-masing kelompok cluster yang dihasilkan. Hasil yang didapatkan dari penelitian tersebut adalah metode clustering dengan algoritma K-Means dapat menghasilkan rekomendasi penerima beasiswa Peningkatan Prestasi Akademik (PPA) dengan melibatkan 7 atribut dan dengan jumlah dataset yang banyak. Penentuan centroid (titik pusat) sangat berpengaruh pada hasil cluster seperti pada hasil pengujian dengan data uji dan data keseluruhan menghasilkan hasil cluster yang berbeda. Pada penelitian yang dilakukan pengujian dengan seluruh data berjumlah 441 dataset, sebanyak 154 mahasiswa direkomendasikan mendapatkan beasiswa PPA sedangkan 287 lainnya tidak mendapatkan, dari 154 data hasil rekomendasi data rill pada tahun 2016 terdapat 113 data pendaftar beasiswa yang dinyatakan diterima dan telah menerima beasiswa PPA, dan setelah dilakukan proses evaluasi oleh biro kemahasiswaan beasiswa tersebut sudah dinyatakan tepat sasaran, sehingga dapat ditarik kesimpulan dengan dilakukan

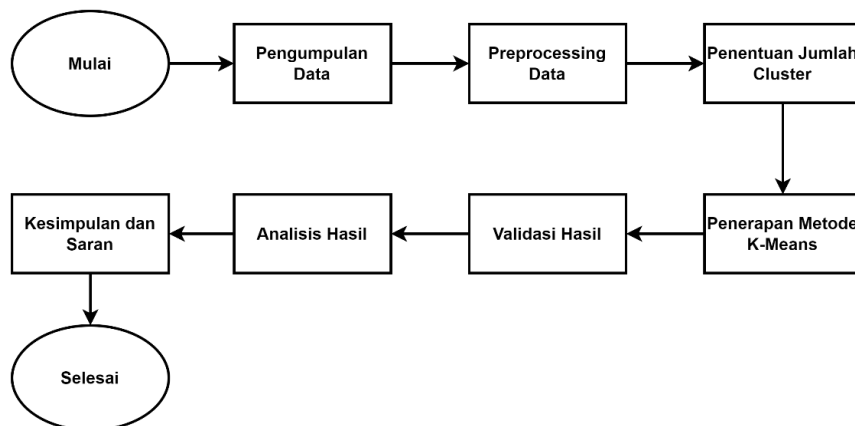
proses rekomendasi pendaftar beasiswa menggunakan metode cluster akan memudahkan dan mempersempit proses seleksi data dengan lebih cepat [16].

Penelitian yang dilakukan oleh Ni Luh Putu Purnama Dewi, I Nyoman Purnama, dan Nengah Widya Utami pada tahun 2022 tentang Penerapan Data Mining Untuk Clustering Penilaian Kinerja Dosen Menggunakan Algoritma K-Means (Studi Kasus: STMIK Primakara). Tujuan dari penelitian tersebut adalah untuk memudahkan memberikan informasi dan evaluasi kepada dosen dan sebagai bahan mengambil keputusan. Metode yang digunakan adalah metode Knowledge Discovery in Database (KDD), yang terdiri dari tahapan: Data Selection, Preprocessing/Cleaning, Transformation Data, Data mining, dan Interpretation/Evaluation. Penerapan metode yang digunakan dalam penelitian ini adalah algoritma K-Means Clustering. Hasil dari penelitian tersebut adalah Algoritma K-Means dapat diimplementasikan pada proses clustering menggunakan tools Rapid Miner. Pengujian cluster dikelompokkan menjadi 4 cluster yaitu: sangat baik, baik, cukup baik, dan kurang baik. Secara garis besar penilaian kinerja dosen oleh mahasiswa didapatkan jumlah cluster penilaian sangat baik 312 (31,74%) data mahasiswa, cluster penilaian baik 401 (40,79%) data mahasiswa, cluster penilaian cukup baik 189 (19,23%) data mahasiswa, dan cluster penilaian kurang baik 81 (8,24%) data mahasiswa. Sehingga jumlah data mahasiswa yang lebih banyak terdapat pada cluster penilaian baik. Hasil evaluasi menggunakan Davies-Bouldin Index diperoleh sebesar 0,270 atau 27%. Sehingga tingkat akurasi dari hasil cluster termasuk baik, karena nilai DBI sudah mendekati nol [17].

Berdasarkan penelitian - penelitian terdahulu tentang metode clustering algoritma K-Means. Dengan menggunakan metode clustering K-Means penelitian ini akan menggunakan data kinerja mahasiswa dari Universitas ABC, termasuk IPK, proyek penelitian, jumlah SKS yang diselesaikan, partisipasi dalam kegiatan akademik dan rekomendasi dari dosen. Data ini kemudian dimasukkan ke dalam algoritma K-Means untuk mengelompokkan mahasiswa ke dalam kelompok prestasi akademik yang berbeda. Selain itu, teknik data mining akan digunakan untuk melakukan analisis lebih lanjut terhadap atribut kinerja mahasiswa dan menemukan pola-pola yang dapat mempengaruhi kelompok prestasi akademik. Penelitian ini akan berfokus pada pengembangan model prediksi kelompok prestasi akademik mahasiswa pada program studi lanjutan menggunakan metode K-Means. Dengan menggabungkan teknik data mining, penelitian ini diharapkan dapat memberikan kontribusi baru dalam memahami dan mengidentifikasi faktor-faktor yang mempengaruhi prestasi akademik mahasiswa pada program studi lanjutan.

3. Metode

Metode pelaksanaan adalah suatu kerangka kerja atau strategi yang digunakan oleh peneliti untuk merancang dan melaksanakan penelitian. Metode pelaksanaan berisi langkah-langkah dan teknik-teknik yang digunakan untuk mengumpulkan dan menganalisis data, serta menjawab pertanyaan penelitian yang telah dirumuskan sebelumnya. Berikut ini adalah metodologi yang digunakan dalam penelitian ini pada gambar 1.



Gambar 1. Tahapan Metodologi Penelitian

- a. Pengumpulan Data
Pertama-tama, melakukan pengumpulan data kinerja akademik mahasiswa dari Universitas ABC. Data ini mencakup informasi tentang nilai mahasiswa pada berbagai mata kuliah, serta beberapa variabel lainnya seperti jenis kelamin, usia, dan fakultas tempat mereka belajar. Data dikumpulkan melalui survei online dan juga dari basis data universitas.
- b. Preprocessing Data
Setelah data terkumpul, peneliti melakukan proses preprocessing untuk membersihkan data dari nilai yang hilang atau tidak valid, serta menghilangkan variabel yang tidak relevan atau berdampak negatif pada analisis. Selain itu, peneliti juga melakukan normalisasi data untuk memastikan bahwa setiap variabel memiliki skala yang sama dan tidak terpengaruh oleh unit ukuran yang berbeda.
- c. Penentuan Jumlah Cluster
Sebelum menerapkan metode K-means, peneliti harus menentukan jumlah cluster yang optimal untuk data. Peneliti menggunakan beberapa metode seperti elbow method dan silhouette method untuk menentukan jumlah cluster yang tepat.
- d. Penerapan Metode K-Means
Setelah menentukan jumlah cluster yang tepat, peneliti menerapkan metode K-means pada data kinerja mahasiswa. Metode ini digunakan untuk mengelompokkan mahasiswa menjadi beberapa kelompok prestasi akademik yang berbeda berdasarkan kinerja mereka.
- e. Validasi Hasil
Pada langkah kali ini akan melakukan validasi hasil untuk memastikan bahwa kelompok prestasi akademik yang dihasilkan oleh metode K-means memiliki perbedaan kinerja yang signifikan dan konsisten dengan karakteristik dan kebutuhan setiap kelompok. Peneliti menggunakan beberapa metode seperti analisis varians (ANOVA) dan uji t untuk memvalidasi hasil.
- f. Analisis Hasil
Setelah hasil divalidasi, kemudian lakukan analisis lebih lanjut terhadap setiap kelompok prestasi akademik yang teridentifikasi. Analisis ini bertujuan untuk mengidentifikasi karakteristik dan kebutuhan setiap kelompok prestasi akademik, serta memberikan rekomendasi untuk pengembangan program pendidikan yang lebih efektif dan efisien di masa yang akan datang.
- g. Kesimpulan dan Saran
Setelah semua Langkah sudah di lakukan langkah terakhir yaitu menyimpulkan hasil penelitian dan memberikan saran untuk pengembangan lebih lanjut di masa depan. Saran ini meliputi penggunaan metode K-means untuk menganalisis data kinerja mahasiswa di universitas lain, serta pengembangan program pendidikan yang lebih terfokus dan efektif berdasarkan karakteristik dan kebutuhan setiap kelompok prestasi akademik. Berdasarkan penelitian - penelitian terdahulu tentang metode clustering algoritma K-Means. Dengan menggunakan metode clustering

4. Hasil dan Pembahasan

a. Analisa Pengelompokan Data

Pada penelitian ini, akan memaparkan hasil dari penerapan metode K-Means pada data kinerja mahasiswa dalam menentukan kelompok prestasi akademik pada program studi lanjutan di Universitas ABC, yang telah diuraikan sebelumnya dalam bab-bab sebelumnya. Dalam penelitian ini, melibatkan sejumlah 20 mahasiswa dari program studi lanjutan di Universitas ABC. Data kinerja mahasiswa yang terdiri dari IPK, proyek penelitian, jumlah SKS yang diselesaikan, partisipasi dalam kegiatan akademik dan rekomendasi dari dosen, telah dikumpulkan dan diproses menggunakan metode K-Means serta teknik data mining untuk mengelompokkan mahasiswa ke dalam kelompok prestasi akademik yang berbeda. Dataset yang diambil dari Universitas ABC akan ditampilkan pada tabel 1.

Table 1. Dataset Mahasiswa

| No. | Nama Mahasiswa | IPK | Jumlah SKS | Proyek Penelitian | Partisipasi Akademik | Rekomendasi Dosen |
|-----|----------------|------|------------|-------------------|----------------------|-------------------|
| D1 | Kiki | 3.75 | 120 | Baik | Tinggi | Baik |
| D2 | Lani | 3.20 | 90 | Cukup | Sedang | Cukup |
| D3 | Mira | 3.90 | 132 | Baik | Tinggi | Sangat Baik |
| D4 | Nanda | 2.85 | 78 | Kurang | Rendah | Kurang |
| D5 | Oki | 3.60 | 105 | Cukup | Sedang | Baik |
| D6 | Putri | 3.95 | 138 | Baik | Tinggi | Sangat Baik |
| D7 | Rudi | 3.80 | 120 | Baik | Tinggi | Baik |
| D8 | Sari | 2.70 | 72 | Kurang | Rendah | Kurang |
| D9 | Toni | 3.50 | 105 | Cukup | Sedang | Baik |
| D10 | Uci | 3.30 | 90 | Cukup | Sedang | Cukup |
| D11 | Vina | 3.85 | 126 | Baik | Tinggi | Sangat Baik |
| D12 | Wulan | 2.65 | 72 | Kurang | Rendah | Kurang |
| D13 | Xena | 3.60 | 120 | Cukup | Tinggi | Baik |
| D14 | Yudi | 3.55 | 111 | Cukup | Sedang | Baik |
| D15 | Zara | 4.00 | 144 | Baik | Tinggi | Sangat Baik |
| D16 | Adit | 2.95 | 78 | Kurang | Rendah | Kurang |
| D17 | Bella | 3.75 | 126 | Baik | Tinggi | Baik |
| D18 | Candra | 3.40 | 99 | Cukup | Sedang | Cukup |
| D19 | Dini | 3.25 | 93 | Kurang | Sedang | Cukup |
| D20 | Evan | 3.95 | 135 | Baik | Tinggi | Sangat Baik |
| D21 | Ali | 3.59 | 120 | Cukup | Tinggi | Baik |
| D22 | Budi | 3.88 | 140 | Baik | Rendah | Sangat Baik |
| D23 | Cici | 3.27 | 110 | Kurang | Rendah | Cukup |
| D24 | Doni | 3.96 | 160 | Baik | Tinggi | Sangat Baik |
| D25 | Eka | 3.65 | 130 | Cukup | Sedang | Baik |
| D26 | Fani | 3.74 | 150 | Baik | Sedang | Sangat Baik |
| D27 | Gilang | 3.43 | 140 | Cukup | Rendah | Cukup |
| D28 | Hani | 3.12 | 100 | Kurang | Sedang | Baik |
| D29 | Indra | 3.81 | 130 | Baik | Tinggi | Sangat |
| D30 | Joko | 3.50 | 120 | Cukup | Sedang | Cukup |

Variabel Yang dipilih pada data kinerja mahasiswa baru yaitu data IPK, jumlah SKS, proyek penelitian, Partisipasi akademik, dan Rekomendasi dosen. Berdasarkan variabel – variabel tersebut terdapat tiga variable yang nilainya masih dalam bentuk kalimat. Berikut merupakan transformasi data dari kalimat kedalam bentuk nilai. Untuk variabel Proyek Penelitian dikelompokkan menjadi 4 kelompok yaitu, Baik ditransformasikan dengan nilai 3, Cukup dengan nilai 2, dan Kurang ditransformasikan dengan nilai 1. Untuk variabel Partisipasi Akademik dikelompokkan menjadi 3 kelompok yaitu, Tinggi ditransformasikan dengan nilai 3, Sedang ditransformasikan menjadi nilai 2 dan Rendah ditransformasikan dengan nilai 1. Untuk variable Rekomendasi Dosen dikelompokkan menjadi 3 kelompok yaitu, Sangat Baik ditransformasikan dengan nilai 4, Baik ditransformasikan dengan nilai 3, Cukup ditransformasikan dengan nilai 2 dan Kurang ditransformasikan dengan nilai 2. Berikut contoh hasil transformasi dapat dilihat pada tabel 2.

Table 2. Data Kinerja Mahasiswa Hasil Tranformasi

| No. | Nama Mahasiswa | IPK | Jumlah SKS | Proyek Penelitian | Partisipasi Akademik | Rekomendasi Dosen |
|-----|----------------|------|------------|-------------------|----------------------|-------------------|
| D1 | Kiki | 3.75 | 120 | 3 | 3 | 3 |
| D2 | Lani | 3.20 | 90 | 2 | 2 | 2 |
| D3 | Mira | 3.90 | 132 | 3 | 3 | 4 |
| D4 | Nanda | 2.85 | 78 | 1 | 1 | 1 |
| D5 | Oki | 3.60 | 105 | 2 | 2 | 3 |

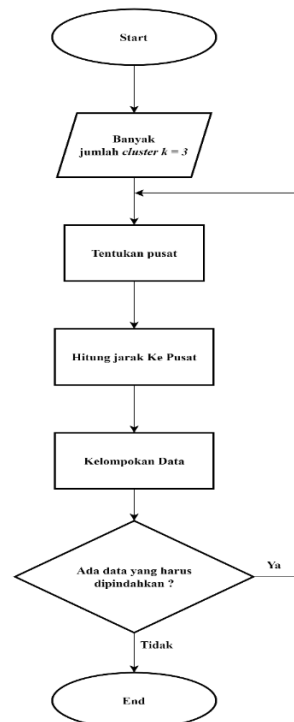


| No. | Nama Mahasiswa | IPK | Jumlah SKS | Proyek Penelitian | Partisipasi Akademik | Rekomendasi Dosen |
|-----|----------------|------|------------|-------------------|----------------------|-------------------|
| D6 | Putri | 3.95 | 138 | 3 | 3 | 4 |
| D7 | Rudi | 3.80 | 120 | 3 | 3 | 3 |
| D8 | Sari | 2.70 | 72 | 1 | 1 | 1 |
| D9 | Toni | 3.50 | 105 | 2 | 2 | 3 |
| D10 | Uci | 3.30 | 90 | 2 | 2 | 2 |
| D11 | Vina | 3.85 | 126 | 3 | 3 | 4 |
| D12 | Wulan | 2.65 | 72 | 1 | 1 | 1 |
| D13 | Xena | 3.60 | 120 | 2 | 3 | 3 |
| D14 | Yudi | 3.55 | 111 | 2 | 2 | 2 |
| D15 | Zara | 4.00 | 144 | 3 | 3 | 4 |
| D16 | Adit | 2.95 | 78 | 1 | 1 | 1 |
| D17 | Bella | 3.75 | 126 | 3 | 3 | 3 |
| D18 | Candra | 3.40 | 99 | 2 | 2 | 2 |
| D19 | Dini | 3.25 | 93 | 1 | 2 | 2 |
| D20 | Evan | 3.95 | 135 | 3 | 3 | 4 |
| D21 | Ali | 3.59 | 120 | 2 | 3 | 3 |
| D22 | Budi | 3.88 | 140 | 3 | 1 | 4 |
| D23 | Cici | 3.27 | 110 | 1 | 1 | 2 |
| D24 | Doni | 3.96 | 160 | 3 | 3 | 4 |
| D25 | Eka | 3.65 | 130 | 2 | 2 | 3 |
| D26 | Fani | 3.74 | 150 | 3 | 2 | 4 |
| D27 | Gilang | 3.43 | 140 | 2 | 1 | 2 |
| D28 | Hani | 3.12 | 100 | 1 | 2 | 3 |
| D29 | Indra | 3.81 | 130 | 3 | 3 | 4 |
| D30 | Joko | 3.50 | 120 | 2 | 2 | 2 |

Tabel dataset di atas memberikan informasi mengenai IPK (Indeks Prestasi Kumulatif) mahasiswa, apakah mereka melaksanakan proyek penelitian atau tidak, jumlah SKS yang telah diselesaikan, tingkat partisipasi dalam kegiatan akademik, dan rekomendasi dari dosen terkait prestasi akademik mereka. Dataset ini dapat digunakan untuk melakukan analisis lebih lanjut menggunakan metode data mining dan eksperimen dalam mengevaluasi faktor-faktor yang mempengaruhi kelompok prestasi akademik mahasiswa pada program studi lanjutan.

b. Perancangan dan Proses Clustering

Pada tahap ini akan melakukan proses utama yaitu segmentasi atau pengelompokan data akademik mahasiswa Universitas ABC dengan menggunakan clustering algoritma K-Means. Berikut diagram alir atau flowchart dari algoritma dengan asumsi banyaknya jumlah cluster $k = 3$ sesuai dengan penelitian.



Gambar 2. Flowchart Proses K-Means

Dari banyak data akademik mahasiswa Universitas ABC yang diperoleh, diambil 30 data untuk dijadikan sampel untuk penerapan algoritma k-means dalam pengelompokan untuk program studi lanjut mahasiswa. Percobaan dilakukan dengan menggunakan parameter-parameter berikut:

- 1) Jumlah *cluster* : 3
- 2) Jumlah data : 30
- 3) Jumlah atribut : 5

Iterasi ke-0

1. Menentukan pusat *cluster* awal.

Menentukan centroid awal dilakukan secara acak dari data/objek yang tersedia sebanyak jumlah cluster k. Nilai *centroid* awal pada penelitian ini dilakukan pemilihan secara acak, di mana jumlah centroid awal dilakukan sebanyak tiga centroid awal, nilai untuk C1 diambil dari baris data ke-15, nilai C2 diambil dari baris ke-21, nilai C3 diambil dari baris data ke-8. Berikut ini nilai centroid awal pada penelitian:

$$C1 = (4.00 \quad ; 144; 3; 3; 4)$$

$$C2 = (3.59 \quad ; 120; 2; 3; 3)$$

$$C3 = (2.70 \quad ; 72; 1; 1; 1)$$

2. Menghitung jarak dengan pusat *cluster*.

Untuk menghitung jarak setiap data yang ada terhadap pusat *cluster* pada penelitian ini penulis menggunakan rumus *Euclidean Distance*.

$$D = \sqrt{[(x_2 - x_1)^2 + (y_2 - y_1)^2]}$$

Di mana,

- (x_1, y_1) adalah koordinat satu titik.
- (x_2, y_2) adalah koordinat titik lainnya.
- d adalah jarak antara (x_1, y_1) dan (x_2, y_2) .

Berikut ini adalah contoh perhitungan jarak dengan *Euclidean Distance* untuk iterasi 0:

Perhitungan jarak data(*objek*) dengan *centroid* 1:

$$D11 = \sqrt{[(3.75 - 4.00)^2 + (120 - 144)^2 + (3 - 3)^2 + (3 - 3)^2 + (3 - 4)^2]} = \sqrt{577.06} = 24.02$$

$$D12 = \sqrt{[(3.20 - 4.00)^2 + (90 - 144)^2 + (2 - 3)^2 + (2 - 3)^2 + (2 - 4)^2]} = \sqrt{2922.64} = 54.06$$

$$D13 = \sqrt{[(3.90 - 4.00)^2 + (132 - 144)^2 + (3 - 3)^2 + (3 - 3)^2 + (4 - 4)^2]} = \sqrt{144.01} = 12.00$$

Perhitungan jarak data(*objek*) dengan *centroid* 2:

$$D21 = \sqrt{[(3.75 - 3.59)^2 + (120 - 120)^2 + (3 - 2)^2 + (3 - 3)^2 + (3 - 3)^2]} = \sqrt{1.03} = 1.01$$

$$D22 = \sqrt{[(3.20 - 3.59)^2 + (90 - 120)^2 + (2 - 2)^2 + (2 - 3)^2 + (2 - 3)^2]} = \sqrt{902.15} = 30.04$$

$$D23 = \sqrt{[(3.90 - 3.59)^2 + (132 - 120)^2 + (3 - 2)^2 + (3 - 3)^2 + (4 - 3)^2]} = \sqrt{146.10} = 12.09$$

Perhitungan jarak data(*objek*) dengan *centroid* 3;

$$D31 = \sqrt{[(3.75 - 2.70)^2 + (120 - 72)^2 + (3 - 1)^2 + (3 - 1)^2 + (3 - 1)^2]} = \sqrt{2317.10} = 48.14$$

$$D32 = \sqrt{[(3.20 - 2.70)^2 + (90 - 72)^2 + (2 - 1)^2 + (2 - 1)^2 + (2 - 1)^2]} = \sqrt{327.25} = 18.09$$

$$D33 = \sqrt{[(3.90 - 2.70)^2 + (132 - 72)^2 + (3 - 1)^2 + (3 - 1)^2 + (4 - 1)^2]} = \sqrt{3618.44} = 60.15$$

3. Mengelompokkan setiap data berdasarkan jarak terdekat antara data dengan *centroid*nya.

Melakukan alokasikan dari masing-masing data ke dalam *centroid* yang paling terdekat. Dalam mengalokasikan kembali data ke dalam masing-masing *cluster* didasarkan pada perbandingan jarak antara data dengan *centroid* setiap *cluster* yang ada, data dialokasikan secara tegas ke dalam *cluster* yang mempunyai jarak ke *centroid* terdekat dengan data tersebut. Berikut ini merupakan hasil perbandingan jarak antara data dengan *centroid* setiap *cluster* yang ada berdasarkan perhitungan jarak dengan Euclidean Distance untuk iterasi 0.

Table 3. Pengelompokan data

| Melakukan Perbandingan | | | | |
|------------------------|----------|----------|----------|-------|
| NO | Jarak C1 | Jarak C2 | Jarak C3 | Class |
| 1 | 24.02 | 1.01 | 48.14 | 2 |
| 2 | 54.06 | 30.04 | 18.09 | 3 |
| 3 | 12.00 | 12.09 | 60.15 | 1 |
| 4 | 66.14 | 42.11 | 6.00 | 3 |
| 5 | 39.04 | 15.03 | 33.10 | 2 |
| 6 | 6.00 | 18.06 | 66.14 | 1 |
| 7 | 24.02 | 1.02 | 48.14 | 2 |
| 8 | 72.13 | 48.10 | 0.00 | 3 |
| 9 | 39.04 | 15.03 | 33.10 | 2 |
| 10 | 54.06 | 30.03 | 18.09 | 3 |
| 11 | 18.00 | 6.17 | 54.17 | 2 |
| 12 | 72.13 | 48.10 | 0.05 | 3 |
| 13 | 24.04 | 0.01 | 48.10 | 2 |
| 14 | 33.09 | 9.11 | 39.05 | 2 |
| 15 | 0.00 | 24.05 | 72.13 | 1 |
| 16 | 66.14 | 42.11 | 6.01 | 3 |
| 17 | 18.03 | 6.08 | 54.12 | 2 |
| 18 | 45.07 | 21.05 | 27.06 | 2 |
| 19 | 51.09 | 27.06 | 21.05 | 3 |
| 20 | 9.00 | 15.07 | 63.15 | 1 |
| 21 | 24.05 | 0.00 | 48.10 | 2 |
| 22 | 4.47 | 20.15 | 68.11 | 1 |
| 23 | 34.18 | 10.30 | 38.02 | 2 |
| 24 | 16.00 | 40.03 | 88.11 | 1 |
| 25 | 14.11 | 10.05 | 58.06 | 2 |
| 26 | 6.09 | 30.05 | 78.10 | 1 |
| 27 | 5.03 | 20.13 | 68.02 | 1 |
| 28 | 44.08 | 20.06 | 28.09 | 2 |
| 29 | 14.00 | 10.10 | 58.16 | 2 |
| 30 | 24.13 | 1.42 | 48.04 | 2 |

Berdasarkan hasil pengelompokan menggunakan titik pusat sementara, didapatkan hasil pengelompokan mahasiswa sebagai berikut:

- Anggota *cluster* 1 (C1): { Mira, Putri, Zara, Evan, Budi, Doni, Fani, Gilang } = 8 orang

- Anggota *cluster 2* (C2): { Kiki, Oki, Rudi, Toni, Vina, Xena, Yudi, Bela, Candara, Ali, Cici, Eka, Hani, Indra, Joko } = 15 orang
 - Anggota *cluster 3* (C3): { Lani, Nanda, Sari, Uci, Wulan, Adit, Dini } = 7 orang
4. Menentukan posisi *centroid* baru (k C) dengan cara menghitung nilai rata – rata dari data – daa yang ada pada *centroid* yang sama. Berikut rumus nya.

$$C_k = \left(\frac{1}{n_k} \right) \sum d_i$$

Dimana,

- k n, adalah jumlah dokumen dalam *cluster k*
- i d adalah dokumen dalam *cluster k*

Menghitung titik pusat baru:

$$C_1 = \left(\begin{array}{c} \frac{3.90 + 3.95 + 4.00 + 3.95 + 3.88 + 3.96 + 3.74 + 3.43}{8}, \\ \frac{132 + 138 + 144 + 135 + 140 + 160 + 150}{8}, \\ \frac{3 + 3 + 3 + 3 + 3 + 3 + 3 + 2}{8}, \\ \frac{3 + 3 + 3 + 3 + 1 + 3 + 2 + 1}{8}, \\ \frac{4 + 4 + 4 + 4 + 4 + 4 + 4 + 2}{8}, \end{array} \right)$$

$$C_1 = \left(\begin{array}{c} 3.85, \\ 142.38, \\ 2.88, \\ 2.38, \\ 3.75 \end{array} \right)$$

Berikut merupakan hasil dari perhitungan titik pusat baru;

Table 4. Tabel Titik Pusat Baru

| Titik Pusat Baru | | |
|------------------|--------|-------|
| C1 | C2 | C3 |
| 3.85 | 3.58 | 2.99 |
| 142.38 | 116.13 | 81.86 |
| 2.88 | 2.20 | 1.29 |
| 2.38 | 2.40 | 1.43 |
| 3.75 | 2.87 | 1.43 |

5. Lakukan iterasi, kemudian tentukan posisi *centroid* baru dengan menggunakan persamaan (1).
6. Mengulangi Langkah ke-3 jika posisi *centroid* baru dengan *centroid* lama tidak sama.

Pengecekan *konvergensi* dilakukan dengan membandingkan matriks *group assignment* pada iterasi sebelumnya dengan matrik *group assignment* pada iterasi yang sedang berjalan. Jika hasilnya sama maka algoritma *kmeans cluster analysis* sudah *konvergen*, tetapi jika berbeda maka belum konvergen sehingga perlu dilakukan iterasi berikutnya. Berikut merupakan hasil dari iterasi 1 sampai iterasi 4.

Table 5. Iterasi 1

| Melakukan Perbandingan Iterasi 1 | | | | |
|----------------------------------|----------|----------|----------|-------|
| NO | Jarak C1 | Jarak C2 | Jarak C3 | Class |
| 1 | 22.40 | 4.00 | 38.25 | 2 |
| 2 | 52.42 | 26.15 | 8.22 | 3 |
| 3 | 10.40 | 15.94 | 50.27 | 1 |
| 4 | 64.48 | 38.23 | 3.92 | 3 |
| 5 | 37.40 | 11.14 | 23.22 | 2 |



| Melakukan Perbandingan Iterasi 1 | | | | |
|----------------------------------|----------|----------|----------|-------|
| NO | Jarak C1 | Jarak C2 | Jarak C3 | Class |
| 6 | 4.43 | 21.92 | 56.26 | 1 |
| 7 | 22.40 | 4.00 | 38.25 | 2 |
| 8 | 70.48 | 44.22 | 9.88 | 3 |
| 9 | 37.40 | 11.14 | 23.22 | 2 |
| 10 | 52.42 | 26.15 | 8.22 | 3 |
| 11 | 16.39 | 9.99 | 44.29 | 2 |
| 12 | 70.48 | 44.22 | 9.89 | 3 |
| 13 | 22.41 | 3.92 | 38.22 | 2 |
| 14 | 31.44 | 5.23 | 29.17 | 2 |
| 15 | 1.77 | 27.91 | 62.25 | 1 |
| 16 | 64.48 | 38.23 | 3.92 | 3 |
| 17 | 16.40 | 9.92 | 44.24 | 2 |
| 18 | 43.42 | 17.16 | 17.18 | 2 |
| 19 | 49.45 | 23.19 | 11.18 | 3 |
| 20 | 7.41 | 18.93 | 53.26 | 1 |
| 21 | 22.41 | 3.92 | 38.22 | 2 |
| 22 | 2.76 | 23.95 | 58.23 | 1 |
| 23 | 32.51 | 6.47 | 28.15 | 2 |
| 24 | 17.64 | 43.89 | 78.23 | 1 |
| 25 | 12.44 | 13.87 | 48.18 | 1 |
| 26 | 7.64 | 33.90 | 68.22 | 1 |
| 27 | 3.40 | 23.92 | 58.15 | 1 |
| 28 | 42.43 | 7.59 | 18.22 | 2 |
| 29 | 12.39 | 13.95 | 48.27 | 1 |
| 30 | 22.47 | 3.99 | 38.16 | 2 |

Table 6. Iterasi 2

| Melakukan Perbandingan Iterasi 2 | | | | |
|----------------------------------|----------|----------|----------|-------|
| NO | Jarak C1 | Jarak C2 | Jarak C3 | Class |
| 1 | 19.92 | 15.29 | 38.25 | 2 |
| 2 | 49.94 | 14.78 | 8.22 | 3 |
| 3 | 7.93 | 27.30 | 50.27 | 1 |
| 4 | 62.01 | 26.87 | 3.92 | 3 |
| 5 | 34.92 | 0.59 | 23.22 | 2 |
| 6 | 2.03 | 33.29 | 56.26 | 1 |
| 7 | 19.92 | 15.30 | 38.25 | 2 |
| 8 | 68.00 | 32.85 | 9.88 | 3 |
| 9 | 34.92 | 0.55 | 23.22 | 2 |
| 10 | 49.94 | 14.78 | 8.22 | 3 |
| 11 | 13.92 | 21.32 | 44.29 | 1 |
| 12 | 68.00 | 32.85 | 9.89 | 3 |
| 13 | 19.94 | 15.26 | 38.22 | 2 |
| 14 | 28.97 | 6.27 | 29.17 | 2 |
| 15 | 4.16 | 39.28 | 62.25 | 1 |
| 16 | 62.01 | 26.87 | 3.92 | 3 |
| 17 | 13.93 | 21.28 | 44.24 | 1 |
| 18 | 40.95 | 5.81 | 17.18 | 2 |
| 19 | 46.97 | 11.83 | 11.18 | 3 |
| 20 | 4.95 | 30.30 | 53.26 | 1 |
| 21 | 19.94 | 15.26 | 38.22 | 2 |
| 22 | 1.45 | 35.30 | 58.23 | 1 |
| 23 | 30.04 | 5.50 | 28.15 | 2 |

| Melakukan Perbandingan Iterasi 2 | | | | |
|---|-----------------|-----------------|-----------------|--------------|
| NO | Jarak C1 | Jarak C2 | Jarak C3 | Class |
| 24 | 20.11 | 55.27 | 78.23 | 1 |
| 25 | 9.97 | 25.24 | 48.18 | 1 |
| 26 | 10.11 | 45.27 | 68.22 | 1 |
| 27 | 2.38 | 35.26 | 58.15 | 1 |
| 28 | 39.95 | 4.90 | 18.22 | 2 |
| 29 | 9.92 | 25.31 | 48.27 | 1 |
| 30 | 20.00 | 15.25 | 38.16 | 2 |

Table 7. Iterasi 3

| Melakukan Perbandingan Iterasi 3 | | | | |
|---|-----------------|-----------------|-----------------|--------------|
| NO | Jarak C1 | Jarak C2 | Jarak C3 | Class |
| 1 | 17.60 | 19.16 | 38.25 | 1 |
| 2 | 47.63 | 10.92 | 8.22 | 3 |
| 3 | 5.62 | 31.17 | 50.27 | 1 |
| 4 | 59.70 | 23.00 | 3.92 | 3 |
| 5 | 32.61 | 4.15 | 23.22 | 2 |
| 6 | 0.76 | 37.16 | 56.26 | 1 |
| 7 | 17.60 | 19.17 | 38.25 | 1 |
| 8 | 65.69 | 28.98 | 9.88 | 3 |
| 9 | 32.61 | 4.14 | 23.22 | 2 |
| 10 | 47.63 | 10.92 | 8.22 | 3 |
| 11 | 11.60 | 25.19 | 44.29 | 1 |
| 12 | 65.69 | 28.98 | 9.89 | 3 |
| 13 | 17.62 | 19.13 | 38.22 | 1 |
| 14 | 26.65 | 10.11 | 29.17 | 2 |
| 15 | 6.45 | 43.15 | 62.25 | 1 |
| 16 | 59.70 | 23.00 | 3.92 | 3 |
| 17 | 11.61 | 25.15 | 44.24 | 1 |
| 18 | 38.63 | 1.98 | 17.18 | 2 |
| 19 | 44.66 | 7.96 | 11.18 | 2 |
| 20 | 2.66 | 34.17 | 53.26 | 1 |
| 21 | 17.62 | 19.13 | 38.22 | 1 |
| 22 | 2.87 | 39.16 | 58.23 | 1 |
| 23 | 27.74 | 9.20 | 28.15 | 2 |
| 24 | 22.43 | 59.13 | 78.23 | 1 |
| 25 | 7.68 | 29.10 | 48.18 | 1 |
| 26 | 12.43 | 49.13 | 68.22 | 1 |
| 27 | 3.42 | 39.11 | 58.15 | 1 |
| 28 | 37.64 | 1.34 | 18.22 | 2 |
| 29 | 7.61 | 29.18 | 48.27 | 1 |
| 30 | 17.69 | 19.10 | 38.16 | 1 |

Table 8. Iterasi 4

| Melakukan Perbandingan Iterasi 4 | | | | |
|---|-----------------|-----------------|-----------------|--------------|
| NO | Jarak C1 | Jarak C2 | Jarak C3 | Class |
| 1 | 12.43 | 16.83 | 40.11 | 1 |
| 2 | 42.45 | 13.30 | 10.07 | 3 |
| 3 | 0.89 | 28.82 | 52.13 | 1 |
| 4 | 54.52 | 25.35 | 2.08 | 3 |
| 5 | 27.43 | 1.88 | 25.08 | 2 |
| 6 | 5.64 | 34.80 | 58.12 | 1 |
| 7 | 12.43 | 16.83 | 40.11 | 1 |
| 8 | 60.51 | 31.34 | 8.02 | 3 |

| Melakukan Perbandingan Iterasi 4 | | | | |
|----------------------------------|----------|----------|----------|-------|
| NO | Jarak C1 | Jarak C2 | Jarak C3 | Class |
| 9 | 27.43 | 1.87 | 25.08 | 2 |
| 10 | 42.45 | 13.30 | 10.07 | 3 |
| 11 | 6.46 | 22.85 | 46.15 | 1 |
| 12 | 60.52 | 31.34 | 8.03 | 3 |
| 13 | 12.45 | 16.77 | 40.08 | 1 |
| 14 | 21.48 | 7.74 | 31.03 | 2 |
| 15 | 11.62 | 40.79 | 64.11 | 1 |
| 16 | 54.52 | 25.35 | 2.08 | 3 |
| 17 | 6.44 | 22.80 | 46.10 | 1 |
| 18 | 33.46 | 4.33 | 19.04 | 2 |
| 19 | 39.48 | 10.31 | 13.04 | 2 |
| 20 | 2.71 | 31.81 | 55.12 | 1 |
| 21 | 12.45 | 16.77 | 40.08 | 1 |
| 22 | 7.78 | 36.79 | 60.09 | 1 |
| 23 | 22.58 | 6.81 | 30.01 | 2 |
| 24 | 27.60 | 56.77 | 80.09 | 1 |
| 25 | 2.62 | 26.73 | 50.04 | 1 |
| 26 | 17.61 | 46.76 | 70.08 | 1 |
| 27 | 7.92 | 36.73 | 60.01 | 1 |
| 28 | 32.47 | 3.40 | 20.08 | 2 |
| 29 | 2.53 | 26.83 | 50.13 | 1 |
| 30 | 12.53 | 16.73 | 40.02 | 1 |

Berdasarkan hasil *clustering* pada program studi lanjutan dari table – table diatas pada program studi lanjutan , Pada hasil iterasi ke-4 posisi *cluster* tidak berubah atau sama dengan posisi *cluster* pada iterasi ke-3, maka dapat disimpulkan bahwa proses iterasi dapat dihentikan pada iterasi ke-4 dengan hasil:

1. Cluster 1 (C1) terdiri dari 17 orang mahasiswa, yaitu Kiki, Mira, Putri, Rudi, Vina, Xena, Zara, Bela, Evan, Ali, Budi, Doni, Eka, Fani, Gilang, Indra, dan Joko. Mahasiswa-mahasiswa dalam cluster ini dapat dikategorikan sebagai "lolos" dalam kelompok prestasi akademik.
2. Cluster 2 (C2) terdiri dari 7 orang mahasiswa, yaitu Oki, Toni, Yudi, Candra, Dini, Cici, dan Hani. Mahasiswa-mahasiswa dalam cluster ini dapat dikategorikan sebagai "uji coba" dalam kelompok prestasi akademik.
3. Cluster 3 (C3) terdiri dari 6 orang mahasiswa, yaitu Lani, Nanda, Sari, Uci, Wulan, dan Adit. Mahasiswa-mahasiswa dalam cluster ini dapat dikategorikan sebagai "tidak lolos" dalam kelompok prestasi akademik.

5. Kesimpulan

Berdasarkan uraian yang telah ada pada bab –bab sebelumnya maka dapat ditarik kesimpulan sebagai berikut:

1. Hasil *clustering* menunjukkan bahwa terdapat perbedaan dalam tingkat prestasi akademik antara kelompok mahasiswa. Cluster 1 menunjukkan kelompok mahasiswa yang berhasil mencapai tingkat prestasi yang lebih tinggi, cluster 2 menunjukkan kelompok yang masih dalam tahap uji coba, sedangkan cluster 3 menunjukkan kelompok yang belum mencapai tingkat prestasi yang diharapkan. Informasi ini dapat memberikan wawasan yang berharga dalam mengidentifikasi faktor-faktor yang mempengaruhi kelompok prestasi akademik mahasiswa pada program studi lanjutan.
2. Metode *Clustering* Algoritma K-Means dapat diterapkan pada pengelompokan Program studi lanjut mahasiswa Universitas ABC, sehingga metode ini sangat membantu pihak akademik dalam menentukan mahasiswa yang lanjut dalam program studi lanjut, mahasiswa dalam tahap uji coba, dan mahasiswa yang tidak lanjut.

3. Hasil pengelompokan data akademik mahasiswa dapat berfungsi sebagai acuan bagi perencana akademik untuk memantau dan mengevaluasi perkembangan kinerja akademik setiap mahasiswa.

Berikut beberapa saran untuk penelitian selanjutnya yang dapat memberikan kontribusi yang lebih besar dalam pengembangan pemahaman tentang kelompok prestasi akademik mahasiswa di Universitas ABC dan bahkan di lingkungan pendidikan yang lebih dalam penelitian ini:

1. Selain kinerja akademik, ada faktor-faktor lain yang dapat mempengaruhi prestasi mahasiswa, seperti faktor sosial, ekonomi, atau kepribadian. Penelitian selanjutnya dapat mencoba mengintegrasikan fitur-fitur ini ke dalam analisis untuk mendapatkan pemahaman yang lebih komprehensif tentang faktor-faktor yang memengaruhi kelompok prestasi akademik.
2. Penelitian selanjutnya diharapkan dapat mempertimbangkan pengelompokan berdasarkan jenis kelamin atau bidang studi. Ini akan membantu memahami apakah ada perbedaan dalam kinerja akademik antara jenis kelamin atau antara mahasiswa di bidang studi yang berbeda. Hal ini dapat memberikan wawasan tentang faktor-faktor yang perlu diperhatikan dalam meningkatkan prestasi akademik dalam konteks tersebut. Penulisan kesimpulan ditulis dengan isi kesimpulan merangkum jawaban dari tujuan penelitian. Kesimpulan berisi kontribusi dari ringkasan hasil penelitian yang diperoleh secara jelas, baik kelebihan, kekurangan serta saran pengembangan mendatang jika terdapat kekurangan dalam penelitian yang dilakukan. Kesimpulan tidak memuat mengenai bahasan teori, opini, harapan serta rencana peneliti kedepannya namun berorientasi kepada data yang dihasilkan. Kesimpulan diharuskan berupa paragraf tidak berupa point atau dituliskan secara *numbering*.

Referensi

- [1] E. D. Sikumbang, "Penerapan Data Mining Penjualan Sepatu Menggunakan Metode Algoritma Apriori," *J. Tek. Komput. AMIK BSI*, vol. Vol 4, No., no. September, pp. 1–4, 2018.
- [2] S. Butsianto and N. T. Mayangwulan, "Penerapan Data Mining Untuk Prediksi Penjualan Mobil Menggunakan Metode K-Means Clustering," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 3, no. 3, pp. 187–201, 2020, doi: 10.32672/jnkti.v3i3.2428.
- [3] H. D. Wijaya and S. Dwiasnati, "Implementasi Data Mining dengan Algoritma Naïve Bayes pada Penjualan Obat," *J. Inform.*, vol. 7, no. 1, pp. 1–7, 2020, doi: 10.31311/ji.v7i1.6203.
- [4] N. K. Surbakti, "Data Mining Pengelompokan Pasien Rawat Inap Peserta BPJS Menggunakan Metode Clustering (Studi Kasus : RSUD. Bangkatan)," *J. Inf. Technol.*, vol. 1, no. 2, pp. 47–53, 2021, doi: 10.32938/jitu.v1i2.1470.
- [5] Yulia and M. Silalahi, "Penerapan Data Mining Clustering Dalam Mengelompokkan Buku Dengan Metode K-Means," *Indones. J. Comput. Sci.*, vol. 10, no. 1, 2021, doi: 10.33022/ijcs.v10i1.3008.
- [6] Luh Putu Wiwien Widhyastuti, I. N. Sukajaya, and Kadek Yota Ernanda Aryanto, "The Customer Profiling berdasarkan Model RFM dengan Metode K-Means pada Institusi Pendidikan untuk menunjang Strategi Bisnis di Masa Pandemi Covid-19," *JTIM J. Teknol. Inf. dan Multimed.*, vol. 4, no. 2, pp. 94–108, 2022, doi: 10.35746/jtim.v4i2.232.
- [7] M. Harahap, Y. Lubis, and Z. Situmorang, "Analisis Pemasaran Bisnis dengan Data Science : Segmentasi Kepribadian Pelanggan berdasarkan Algoritma K-Means Clustering," *Data Sci. Indones.*, vol. 1, no. 2, pp. 76–88, 2022, doi: 10.47709/dsi.v1i2.1348.
- [8] A. S. Soetoko and D. F. Soetoko, "Meningkatkan Kualitas Hidup Penderita Gangguan Metabolik di Klinik Pratama Semarang melalui Pelatihan Pengaturan Diet dan Senam Kaki," *J. ABDIMAS-KU J. Pengabd. Masy. Kedokt.*, vol. 1, no. 3, p. 126, 2022, doi: 10.30659/abdimasku.1.3.126-133.
- [9] J. Wandana, S. Defit, and S. Sumijan, "Klasterisasi Data Rekam Medis Pasien Pengguna Layanan BPJS Kesehatan Menggunakan Metode K-Means," *J. Inf. dan Teknol.*, vol. 2, pp. 4–9, 2020, doi: 10.37034/jidt.v2i4.73.
- [10] S. Naning, R. Pranata, A. N. M. Clara, and O. S. Rahajeng, "Pengelompokan Kunjungan Wisata Kabupaten Kulon Progo Tahun 2019 Menggunakan K-Means Clustering," *JUST IT J. Sist. ...*, vol. 12, no. 1, pp. 38–45, 2022, [Online]. Available: <https://jurnal.umj.ac.id/index.php/just-it/article/view/8249%0Ahttps://jurnal.umj.ac.id/index.php/just-it/article/download/8249/7242>.
- [11] B. Setio and P. Prasetyaningrum, "Penerapan Data Mining Dalam Mengelompokkan Kunjungan Wisatawan Di Kota Yogyakarta Menggunakan Metode K-Means," *J. Comput. Sci. Technol.*, vol. 1, no. 1, pp. 27–32, 2021, doi: 10.54840/jcstech.v1i1.9.

- [12] Rahmat Hidayat, "Pemanfaatan Data Mining Untuk Melihat Minat Siswa Setelah Menyelesaikan Pendidikan Sekolah Menengah Atas (Sma) Dengan Algoritma K-Means Clustering," *Technol. Informatics Insight J.*, vol. 1, no. 2, pp. 85–97, 2022, doi: 10.32639/tij.v1i2.220.
- [13] A. E. Wibowo and T. Habanabakize, "K-Means Clustering untuk Klasifikasi Standar Kualifikasi Pendidikan dan Pengalaman Kerja Guru SMK di Indonesia," *J. Din. Vokasional Tek. Mesin*, vol. 7, no. 2, pp. 152–163, 2022, doi: 10.21831/dinamika.v7i2.53848.
- [14] E. S. Pujiharti, "Peran Sumber Daya Pendidik Dalam Perspektif Ekonomi Pendidikan," *J. Manaj. Pendidik. Islam*, vol. 1, no. 2, pp. 35–50, 2022.
- [15] H. Priyatman, F. Sajid, and D. Haldivany, "Klasterisasi Menggunakan Algoritma K-Means Clustering untuk Memprediksi Waktu Kelulusan Mahasiswa," *J. Edukasi dan Penelit. Inform.*, vol. 5, no. 1, p. 62, 2019, doi: 10.26418/jp.v5i1.29611.
- [16] A. Salam, D. Adiatma, and J. Zeniarja, "Implementasi Algoritma K-Means Dalam Pengklasteran untuk Rekomendasi Penerima Beasiswa PPA di UDINUS," *JOINS (Journal Inf. Syst.*, vol. 5, no. 1, pp. 62–68, 2020, doi: 10.33633/joins.v5i1.3350.
- [17] N. L. P. P. Dewi, I. N. Purnama, and N. W. Utami, "Penerapan Data Mining Untuk Clustering Penilaian Kinerja Dosen Menggunakan Algoritma K-Means (Studi Kasus: STMIK Primakara)," *J. Ilm. Teknol. Inf. Asia*, vol. 16, no. 2, p. 105, 2022, doi: 10.32815/jitika.v16i2.761.