

Penerapan Metode K-Means Clustering Dalam Pengembangan Strategi Promosi Berbasis Data Penerimaan Mahasiswa Baru (Studi Kasus :Universitas Nurul Huda)

**Indra Irawan¹, Uli Rizki², Pamuji M. Jakak³,
M. Bagus Prayogi⁴, Miftakhul Rahman⁵**

^{1,2,3,4,5}Informatika, Universitas Nurul Huda

Penulis Korespondensi: Indrairawan@unuha.ac.id^{1*},
uli@unuha.ac.id², jakak@unuha.ac.id³, mhdjesen212@gmail.com⁴,
miftakhul.rahman@gmail.com⁵

Informasi Artikel:

Received: 02 Februari 2024, Accepted: 9 Februari 2024, Published: 12 Februari 2024

Abstract

The admission of new students is a key element in the success of a university. In its effort to enhance efficiency in admitting new students, Nurul Huda University proposes applying the K-Means Clustering method as a solution to manage candidate student data more effectively and intelligently. The data used in this research is derived from the Admission of New Students process for the academic years 2021/2022 and 2022/2023, totaling 1275 entries, which will then be processed using data mining techniques to generate analysis. The goal is to determine promotional strategies based on the origin of profiles of newly enrolled students. The method applied is clustering with the K-Means algorithm. After data processing, analysis is conducted using the Knowledge Discovery in Databases (KDD) technique, consisting of five stages: selection, preprocessing, transformation, data mining, and evaluation. The implementation in this research utilizes Rapidminer software, resulting in three data clusters: Cluster 1 with 345 entries, covering 27% of the total; Cluster 2 with 86 entries (7%); and Cluster 3 with 835 entries (66%). For promotion, the marketing team is deployed to districts dominant in the East OKU region and potential areas outside the East OKU District. They conduct direct visits to introduce Nurul University to students, distribute brochures, display pamphlets, and adapt strategies using a promotion mix strategy.

Keywords: *Data Mining, K-Means Clustering, New Student Admission, Promotion Strategi, RapidMiner*

1. Pendahuluan

1.1 Latar Belakang

Dalam konteks pendidikan, terutama di perguruan tinggi, penerimaan mahasiswa baru adalah hal yang krusial dan rutin dilakukan setiap tahun. Pendidikan sendiri adalah proses yang bertujuan untuk mengembangkan karakteristik dan memastikan kualitas [1]. Perguruan tinggi berperan sebagai lembaga yang mendorong peningkatan kualitas dengan standar yang harus terpenuhi sesuai tujuan pendidikan [1]. Universitas Nurul Huda adalah institusi pendidikan tinggi swasta dengan 3 Fakultas diantaranya Fakultas Ilmu Pendidikan, Fakultas Agama Islam dan Fakultas Sains dan Teknologi dengan 10 program studi yaitu S1-Pendidikan Bahasa dan Sastra Indonesia, S1-Pendidikan Bahasa Inggris, S1-Pendidikan Ekonomi, S1-Pendidikan Fisika, S1-Pendidikan Teknologi Informasi, S1-Pendidikan Guru Madrasah Ibtidaiyah, S1-Pendidikan Agama Islam, S1-Informatika, S1-Pertanian dan S1-Matematika. Dalam proses penerimaan mahasiswa baru, strategi promosi diperlukan untuk menarik calon mahasiswa dari dalam dan luar wilayah.

Penerimaan Mahasiswa Baru menghasilkan banyak data tentang profil mahasiswa yang mendaftar. Universitas Nurul Huda telah menggunakan berbagai cara promosi, seperti pemasangan baliho, distribusi selebaran, pemberian potongan biaya, dan kunjungan ke SMA/SMK di Kabupaten OKU Timur. Melalui analisis data yang tepat, informasi berharga dari profil mahasiswa dapat ditemukan dan digunakan untuk mengembangkan strategi promosi yang lebih efisien berdasarkan kelompok daerah asal mahasiswa.

Penelitian sebelumnya telah menggunakan metode K-Means Clustering untuk mengelompokkan calon penerima beasiswa menjadi 3 kelompok [2]. Penelitian lainnya mengidentifikasi strategi promosi berdasarkan asal sekolah dan jurusan yang diambil dengan menggunakan 2 kelompok [3]. Metode K-Means Clustering juga digunakan untuk membantu pihak marketing dalam promosi berdasarkan asal mahasiswa yang lulus pada tahun 2012 [4]. Penelitian lainnya dengan judul Segmentasi Lokasi Promosi Penerimaan Mahasiswa Baru Menggunakan Metode RFM dan K-Means Clustering dimana hasil penelitian adalah terbentuk 3 klaster tingkat potensial sekolah yang dapat dijadikan skala prioritas untuk lokasi promosi penerimaan mahasiswa baru Universitas Bumigora [1]. Peneliti tertarik untuk menerapkan metode K-Means Clustering dalam mengelompokkan daerah asal mahasiswa baru di Universitas Nurul Huda. Software Rapidminer digunakan untuk mempermudah dalam analisis. Judul dari penelitian ini adalah ' Peningkatan Efisiensi Penerimaan Mahasiswa Baru Universitas Nurul Huda Melalui Penerapan Metode K-Means Clustering '.

Masalah yang perlu dipecahkan dalam penelitian ini adalah bagaimana metode K-Means Clustering dapat diterapkan untuk dapat meningkatkan efisiensi dalam menentukan strategi promosi berdasarkan data penerimaan mahasiswa baru di Universitas Nurul Huda.

2. Tinjauan Literatur

2.1 Data Mining

Data mining adalah hasil perpaduan antara beberapa bidang ilmu yang menggabungkan teknik-teknik dari machine learning, pengenalan pola, statistik, manajemen basis data, serta visualisasi. Tujuan utamanya adalah mengatasi permasalahan dalam mengambil

informasi berharga dari kumpulan data besar dalam sebuah database yang signifikan [2].

2.2 Clustering

Clustering adalah proses mengelompokkan objek-objek serupa ke dalam kelompok yang berbeda, atau lebih tepatnya membagi sebuah set data menjadi subset-subset, sehingga setiap subset memiliki makna yang berguna. Setiap kelompok, yang disebut cluster, terdiri dari objek-objek yang memiliki kesamaan satu sama lain dan berbeda dari objek-objek yang ada di cluster lainnya [3].

Clustering juga dapat digambarkan sebagai suatu proses di mana data diorganisir dan dibagi menjadi beberapa kelompok sehingga membentuk pola-pola serupa yang dikelompokkan bersama dalam cluster yang sama, sementara pola yang berbeda ditempatkan dalam cluster yang berbeda[6].

Clustering dapat memainkan peran penting dalam kehidupan sehari-hari, karena tidak bisa lepas dengan sejumlah data yang menghasilkan informasi untuk memenuhi kebutuhan hidup. Salah satu sarana yang paling penting dalam hubungan dengan data adalah untuk mengklasifikasikan atau mengelompokkan data tersebut ke dalam seperangkat kategori atau cluster. Clustering dapat ditemukan di beberapa aplikasi yang ada di berbagai bidang. Sebagai contoh pengelompokan data yang digunakan untuk menganalisa data statistik seperti pengelompokan untuk pembelajaran mesin, data mining, pengenalan pola, analisis citra dan bioinformatika[6].

2.3 Algoritma K-Means

Algoritma K-Means adalah salah satu algoritma partitional yang berdasarkan pada penentuan jumlah awal kelompok dengan menetapkan nilai awal centroid [7]. Algoritma K-Means menggunakan proses berulang untuk membentuk kelompok data. Jumlah cluster awal yang diinginkan dimasukkan sebagai input, dan algoritma ini menghasilkan titik centroid akhir sebagai output. Metode K-Means akan memilih titik awal centroid secara acak. Jumlah iterasi yang diperlukan untuk mencapai centroid cluster dipengaruhi oleh inisialisasi awal yang acak. Oleh karena itu, ada upaya untuk mengembangkan algoritma dengan menentukan centroid cluster yang berdasarkan pada kepadatan data awal yang tinggi, untuk meningkatkan kinerjanya [7]. Dasar algoritma K-means adalah sebagai berikut :

1. Tentukan nilai k sebagai jumlah klaster yang ingin dibentuk.
2. Inisialisasi k sebagai centroid yang dapat dibangkitkan secara random.
3. Hitung jarak setiap data ke masing-masing centroid menggunakan persamaan Euclidean Distance yaitu sebagai berikut :
4. Kelompokkan setiap data berdasarkan jarak terdekat antara data dengan centroidnya.

$$d(P,Q) = \sqrt{\sum_{j=1}^p (x_j(P) - x_j(Q))^2}$$

5. Tentukan posisi centroid baru (k)
6. kembali ke langkah 3 jika posisi centroid baru dengan centroid lama tidak sama.

2.4 RapidMiner

RapidMiner ialah perangkat lunak yang dirancang untuk mengolah data dengan menerapkan prinsip dan algoritma data mining. Dengan menggunakan kombinasi metode statistika, kecerdasan buatan, dan database, RapidMiner mengekstrak pola-pola dari kumpulan data yang besar. Pengguna dapat dengan mudah melakukan perhitungan pada volume data yang besar menggunakan operator-operator yang disediakan[8]. Operator tersebut berfungsi untuk memodifikasi data, dengan menghubungkan data ke node-node pada operator. Selanjutnya, hasil perhitungan dapat divisualisasikan dalam bentuk grafik dengan menghubungkannya ke node hasil. RapidMiner menyajikan hasil secara visual, memudahkan pengguna untuk memahami dan menganalisis informasi yang diekstrak [8].

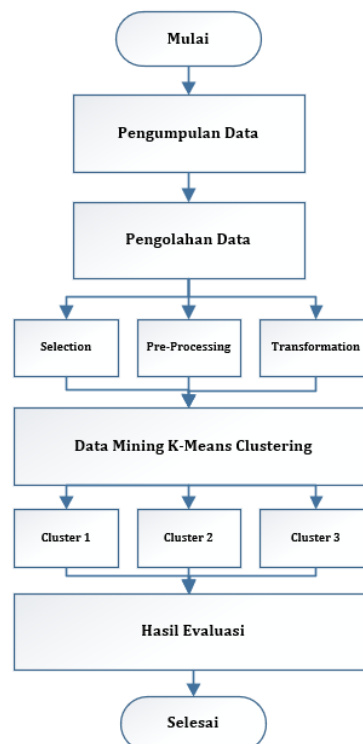
2.5 Promosi

Promosi adalah tindakan yang dilakukan untuk mengembangkan komunikasi antara institusi dan calon mahasiswa baru beserta keluarganya, dengan maksud memengaruhi keputusan mereka dalam memilih lembaga pendidikan[9].

3. Metode Penelitian

3.1 Tahapan Penelitian

Untuk mendapatkan hasil dari penelitian ini, peneliti menyajikan sebuah flowchart yang dapat dilihat pada Gambar 1.



Gambar 1. Tahapan Penelitian

3.1.1 Tahap Pengumpulan Data

Dalam tahap ini peneliti melakukan pengumpulan data di Universitas Nurul Huda. Data yang dikumpulkan berupa data dari proses Penerimaan Mahasiswa Baru Tahun Ajaran 2021/2022 dan 2022/2023.

3.1.2 Tahap pengolahan Data

Data mentah yang terkumpul akan mengalami tahap pengeditan, termasuk langkah-langkah seperti menghilangkan entri kosong, melengkapi informasi yang kurang jelas, dan memperbaiki kesalahan yang mungkin ada dalam data tersebut.

3.1.2 Tahap Analisa Data

Metode yang diterapkan dalam proses analisis adalah *Knowledge Discovery in Databases* (KDD), yang merupakan suatu proses non-trivial untuk menemukan pola-pola baru yang bermanfaat dan mudah dimengerti [6]. Proses ini melibatkan lima langkah berurutan, yaitu pemilihan, pre-processing, transformasi, eksplorasi data, dan evaluasi [5].

a. Selection

Seleksi data adalah langkah awal di mana data diambil dan dianalisis secara cermat untuk kemudian dapat diproses dalam analisis data mining. Hasil dari proses seleksi data dapat ditemukan dalam Tabel 1.

Tabel 1 Sampel Hasil Seleksi Data

No	Jenis Kelamin	Asal Sekolah	Alamat Mahasiswa	Program Studi
1	Perempuan	SMA Terpadu Ponpes Nurul huda 2	Pekuncen kel. gunung mas, Kec. Belitang 1, Kab. OKU Timur	S1 - Pendidikan Agama Islam
2	Laki-laki	SMK Darulssalam	Taman Harjo Kec.Se-mendawai Suku III Kab.OKU Timur	S1 - Informatika
3	Perempuan	MA Tajul Ulum	Sri Bulan RT.003 RW.001 kecamatan Buay Pemuka Bangsa Raja Kab.OKU Timur	S1 - Pendidikan Agama Islam
4	Laki – laki	MA Maarif NU 02 Sidorejo Lampung Timur	jl. Wana Sari Kel.Wana Sari Kec.Semendawai Timur, Kab.OKU Timur	S1 – Pendidikan Teknologi Informasi
5	Perempuan	SMA N 1 Buay Madang	Desa muncak kabau Kel.Muncak Kabau, Inggris Kec.Buay Pemuka Bangsa Raja, Kab.OKU Timur	S1 – Pendidikan Bahasa Inggris

b. Pre-processing

Pu

Pre-processing tahap untuk menghilangkan noise, mendeteksi dan mengatasi duplikasi, serta menangani ketidak konsistenan dan kekosongan data pada set data yang tersedia.

c. *Transformations*

Proses ini bertujuan untuk melakukan perubahan dan konversi pada data yang belum memiliki entitas yang jelas.

1. Transformasi

Atribut yang akan mengalami transformasi adalah sebagai berikut:

a. Atribut Asal Sekolah

Data pada atribut Asal Sekolah awalnya berupa identitas sekolah sebelumnya. Dari beragam sekolah tersebut, kemudian dikelompokkan menjadi empat kategori: "SMA", "SMK", "MA", "Lainya". Informasi lebih lanjut dapat ditemukan dalam Tabel 2 di bawah ini.

Tabel 2. Transformasi Atribut Asal Sekolah

No	Asal Sekolah	Hasil Transformasi
1	SMA Terpadu Ponpes Nurul Huda 2	SMA
2	SMA N 1 Buay Madang	
3	SMK Darulssalam	SMK
4	MA Maarif NU 02 Sidorejo Lampung Timur	MA
5	MA Tajul Ulum	

b. Atribut Alamat Mahasiswa

Informasi pada atribut Alamat Mahasiswa mencakup daerah berdasarkan kecamatan dan Kabupaten tempat tinggal masing-masing mahasiswa. Berikut adalah tabel transformasi alamat mahasiswa. Berikut sample transformasi alamat mahasiswa dapat dilihat pada tabel 3.

Tabel 3. Transformasi Alamat Mahasiswa

Alamat Mahasiswa (Kabupaten)	Hasil Konversi (Kecamatan)
OKU TIMUR	Kec. Belitang I
	Kec. Belitang II
	Kec. Belitang III
	Kec. Belitang Jaya
	Kec. Belitang Madang Raya
	Kec. Belitang Mulia
	Kec. Buay Madang
	Kec. Buay Madang Timur
	Kec. Buay Pemuka Bangsa Raja
	Kec. Buay Pemuka Peliung
	Kec. Bunga Mayang
	Kec. Cempaka
	Kec. Jayapura
	Kec. Madang Suku I
	Kec. Madang Suku II
	Kec. Madang Suku III
	Kec. Martapura
	Kec. Semendawai Barat
	Kec. Semendawai Suku III
	Kec. Semendawai Timur
WAY KANAN	Kec. Buay Bahuga
MUARA ENIM	Kec. Gelumbang

2. Konversi

Konversi Data yang sudah di transformasi kemudian dilakukan proses konversi dengan tujuan untuk mempermudah dalam proses data mining. Dalam konversi data semua atribut akan di ubah menjadi angka. Hasil data dari keseluruhan proses pre-processing dapat dilihat pada sampel data berikut .

Tabel 4. Data Sampel Hasil Pre-processing

Jenis Kelamin	Asal Sekolah	Alamat Mahasiswa	Program Studi
2	1	1	1
2	3	8	7
1	2	2	1
1	2	1	2
1	4	3	4

d. Data Mining

Data Mining merupakan suatu proses untuk mengidentifikasi pola atau informasi menarik dari suatu kumpulan data menggunakan teknik, metode, atau algoritma tertentu[2]. Proses ini menghasilkan pengetahuan baru yang berasal dari data yang sudah ada, yang nantinya dapat digunakan dalam pengambilan keputusan di masa depan[6]. Salah satu tahap penting dalam membentuk pola dari data tersebut adalah menggunakan metode clustering. Clustering adalah algoritma yang digunakan untuk mengelompokkan data yang berjumlah besar menjadi bagian-bagian kecil berdasarkan atribut kemiripan tertentu[5]. Analisis clustering dilakukan dengan membagi satu set data menjadi beberapa bagian yang memiliki kesamaan dalam satu kelompok, dan kelompok tersebut memiliki perbedaan karakteristik dengan kelompok lainnya[10].

Dalam penelitian ini, metode yang digunakan untuk melakukan clustering adalah algoritma K-Means Clustering. K-Means Clustering merupakan salah satu metode dalam proses data mining yang berfokus pada pengelompokkan data dengan karakteristik yang serupa antara satu sama lain, sementara memiliki perbedaan karakteristik dengan data dari kelompok lain[3]. Perhitungan klasterisasi menggunakan perangkat lunak *Rapidminer* yang akan menghasilkan tiga *cluster* dengan jumlah record yang berbeda, namun dengan variabel yang serupa. Pengolahan data mahasiswa baru dilakukan setelah proses transformasi sehingga data mahasiswa baru bisa diolah menggunakan metode K-Mean Clustering. Adapun langkah-langkah proses algoritma K-Mean Clustering adalah sebagai berikut:

1. Dilakukan k dari jumlah cluster baru yang ingin dibentuk. Adapun cluster yang

akan dibuat adalah 3 cluster.

2. Tentukan titik pusat awal dari setiap cluster. Adapun penentuan titik pusat awal dalam penelitian ini ditentukan secara acak dan titik pusat yang didapatkan dapat dilihat pada tabel berikut:

Tabel 5. Titik Pusat awal

Atribut	Cluster 1	Cluster 2	Cluster 3
Jenis Kelamin	2	1	1
Asal Sekolah	1	2	2
Alamat Mahasiswa	1	1	2
Program Studi	1	2	1

3. Menghitung jarak dari setiap data ke pusat klaster yang sudah ditentukan dengan menggunakan persamaan rumus Eucliden Distance, sebagai berikut:

$$d(P, Q) = \sqrt{\left(\sum_{j=1}^p (x_j(P) - x_j(Q))^2\right)}$$

Keterangan:

d = titik dokumen

P = data record

Q = data centroid

Contoh perhitungan Data Cluster 1 dapat dilihat sebagai berikut:

$$d(1,1) = \sqrt{(X_1 - Y_1)^2 + (X_2 - Y_2)^2 + (X_3 - Y_3)^2 + (X_4 - Y_4)^2}$$

$$d(1,1) = \sqrt{(2 - 2)^2 + (1 - 1)^2 + (1 - 1)^2 + (1 - 1)^2}$$

$$d(1,1) = \sqrt{1}$$

$$d(1,1) = 1$$

$$d(1,2) = \sqrt{(X_1 - Y_1)^2 + (X_2 - Y_2)^2 + (X_3 - Y_3)^2 + (X_4 - Y_4)^2}$$

$$d(1,2) = \sqrt{(2 - 2)^2 + (3 - 1)^2 + (8 - 1)^2 + (7 - 1)^2}$$

$$d(1,2) = \sqrt{89}$$

$$d(1,2) = 9,434$$

$$d(1,3) = \sqrt{(X_1 - Y_1)^2 + (X_2 - Y_2)^2 + (X_3 - Y_3)^2 + (X_4 - Y_4)^2}$$

$$d(1,3) = \sqrt{(1 - 2)^2 + (2 - 1)^2 + (2 - 1)^2 + (1 - 1)^2}$$

$$d(1,3) = \sqrt{1}$$

$$d(1,3) = 1$$

$$d(1,4) = \sqrt{X_1 - Y_1)^2 + (X_2 - Y_2)^2 + (X_3 - Y_3)^2 + (X_4 - Y_4)^2}$$

$$d(1,4) = \sqrt{(1 - 2)^2 + (2 - 1)^2 + (1 - 1)^2 + (2 - 1)^2}$$

$$d(1,4) = \sqrt{0}$$

$$d(1,4) = 0$$

$$d(1,5) = \sqrt{X_1 - Y_1)^2 + (X_2 - Y_2)^2 + (X_3 - Y_3)^2 + (X_4 - Y_4)^2}$$

$$d(1,5) = \sqrt{(1 - 2)^2 + (4 - 1)^2 + (3 - 1)^2 + (4 - 1)^2}$$

$$d(1,5) = \sqrt{21}$$

$$d(1,5) = 4,582$$

Contoh perhitungan Data Cluster 2 dapat dilihat sebagai berikut:

$$d(1,1) = \sqrt{(X_1 - Y_1)^2 + (X_2 - Y_2)^2 + (X_3 - Y_3)^2 + (X_4 - Y_4)^2}$$

$$d(1,1) = \sqrt{(2 - 1)^2 + (1 - 2)^2 + (1 - 1)^2 + (1 - 2)^2}$$

$$d(1,1) = \sqrt{0}$$

$$d(1,1) = 0$$

$$d(1,2) = \sqrt{(X_1 - Y_1)^2 + (X_2 - Y_2)^2 + (X_3 - Y_3)^2 + (X_4 - Y_4)^2}$$

$$d(1,2) = \sqrt{(2 - 1)^2 + (3 - 2)^2 + (8 - 1)^2 + (7 - 2)^2}$$

$$d(1,2) = \sqrt{73}$$

$$d(1,2) = 8,44$$

$$d(1,3) = \sqrt{(X_1 - Y_1)^2 + (X_2 - Y_2)^2 + (X_3 - Y_3)^2 + (X_4 - Y_4)^2}$$

$$d(1,3) = \sqrt{(1 - 1)^2 + (2 - 2)^2 + (2 - 1)^2 + (1 - 2)^2}$$

$$d(1,3) = \sqrt{0}$$

$$d(1,3) = 0$$

$$d(1,4) = \sqrt{(X_1 - Y_1)^2 + (X_2 - Y_2)^2 + (X_3 - Y_3)^2 + (X_4 - Y_4)^2}$$

$$d(1,4) = \sqrt{(1 - 1)^2 + (2 - 2)^2 + (1 - 1)^2 + (2 - 2)^2}$$

$$d(1,4) = \sqrt{1}$$

$$d(1,4) = 1$$

$$d(1,5) = \sqrt{(X_1 - Y_1)^2 + (X_2 - Y_2)^2 + (X_3 - Y_3)^2 + (X_4 - Y_4)^2}$$

$$d(1,5) = \sqrt{(1 - 1)^2 + (4 - 2)^2 + (3 - 1)^2 + (4 - 2)^2}$$

$$d(1,5) = \sqrt{12}$$

$$d(1,5) = 3,464$$

Contoh perhitungan Data Cluster 3 dapat dilihat sebagai berikut:

$$d(1,1) = \sqrt{(X_1 - Y_1)^2 + (X_2 - Y_2)^2 + (X_3 - Y_3)^2 + (X_4 - Y_4)^2}$$

$$d(1,1) = \sqrt{(2 - 1)^2 + (1 - 2)^2 + (1 - 2)^2 + (1 - 1)^2}$$

$$d(1,1) = \sqrt{1}$$

$$d(1,1) = 1$$

$$d(1,2) = \sqrt{(X_1 - Y_1)^2 + (X_2 - Y_2)^2 + (X_3 - Y_3)^2 + (X_4 - Y_4)^2}$$

$$d(1,2) = \sqrt{(2-1)^2 + (3-2)^2 + (8-2)^2 + (7-1)^2}$$

$$d(1,2) = \sqrt{74}$$

$$d(1,2) = 8,602$$

$$d(1,3) = \sqrt{(X_1 - Y_1)^2 + (X_2 - Y_2)^2 + (X_3 - Y_3)^2 + (X_4 - Y_4)^2}$$

$$d(1,3) = \sqrt{(1-1)^2 + (2-2)^2 + (2-2)^2 + (1-1)^2}$$

$$d(1,3) = \sqrt{0}$$

$$d(1,3) = 0$$

$$d(1,4) = \sqrt{(X_1 - Y_1)^2 + (X_2 - Y_2)^2 + (X_3 - Y_3)^2 + (X_4 - Y_4)^2}$$

$$d(1,4) = \sqrt{(1-1)^2 + (2-2)^2 + (1-2)^2 + (2-1)^2}$$

$$d(1,4) = \sqrt{1}$$

$$d(1,4) = 1$$

$$d(1,5) = \sqrt{(X_1 - Y_1)^2 + (X_2 - Y_2)^2 + (X_3 - Y_3)^2 + (X_4 - Y_4)^2}$$

$$d(1,5) = \sqrt{(1-1)^2 + (4-2)^2 + (3-2)^2 + (4-1)^2}$$

$$d(1,5) = \sqrt{14}$$

$$d(1,5) = 3,741$$

Tabel 6. Hasil Perhitungan Jarak terdekat

Data ke- i	Jenis Kelamin	Asal Sekolah	Alamat	Program Studi	C1	C2	C3	Cluster
1	2	1	1	1	1	0	1	C2
2	2	3	8	7	9,434	8,602	8,544	C3
3	1	2	2	1	1	0	1	C2
4	1	2	1	2	0	1	1	C1
5	1	4	3	4	4,582	3,464	3,741	C2

4. Tentukan posisi centroid baru (k) dengan rumus berikut:

$$ClusterCenter = \sum \frac{a_i}{n}$$

Berikut contoh perhitungan cluster center iterasi ke -2.

$$C1 = \text{data ke- 4} \rightarrow (1,2,1,2)$$

$$C2 = \text{data ke- 1, 3, 5} \rightarrow (2,1,1,1), (1,2,2,1), (1,4,3,4)$$

$$C3 = \text{data ke- 2} \rightarrow (2,3,8,7)$$

Masukan kedalam rumus *cluster center* pada tabel berikut

Tabel 7. Perhitungan *cluster center*

Cluster 1	Cluster 2	Cluster 3
Jenis Kelamin $\rightarrow 1/1 = 1$	Jenis Kelamin $\rightarrow 2+1+1/3 = 1,333$	Jenis Kelamin $\rightarrow 2/1 = 2$
Asal Sekolah $\rightarrow 2/1 = 2$	Asal Sekolah $\rightarrow 1+2+4/3 = 2,333$	Asal Sekolah $\rightarrow 3/1 = 3$
Alamat $\rightarrow 1/1 = 1$	Alamat $\rightarrow 1+2+3/3 = 2$	Alamat $\rightarrow 8/1 = 8$
Program Studi $\rightarrow 2/1 = 2$	Program Studi $\rightarrow 1+1+4/3 = 2$	Program Studi $\rightarrow 7/1 = 7$

Sehingga didapatkan nilai cluster center untuk iterasi ke 2 sebagai berikut :

Tabel 8. Cluster center iterasi ke-2

Atribut	Cluster 1	Cluster 2	Cluster 3
Jenis Kelamin	1	1,333	2
Asal Sekolah	2	2,333	3
Alamat Mahasiswa	1	2	8
Program Studi	1	2	7

Lakukan kembali langkah nomor 3 jika pusat klaster masih berubah sampai pusat klaster tidak berubah lagi. Hasil Akhir Perhitungan Cluster dapat dilihat pada Tabel 11.

Tabel 9. Hasil Akhir Perhitungan Cluster

Data ke- i	Jenis Kelamin	Asal Sekolah	Alamat	Progrm Studi	Hasil Cluster Iterasi ke-1	Hasil Cluster Iterasi ke-2	Hasil Cluster Iterasi ke-3	Hasil Cluster Iterasi ke-4
1	2	1	1	1	C2	C2	C2	C2
2	2	3	8	7	C3	C2	C3	C1
3	1	2	2	1	C2	C3	C1	C1
4	1	2	1	2	C1	C1	C1	C1
5	1	4	3	4	C2	C2	C2	C3

e. Hasil Evaluasi

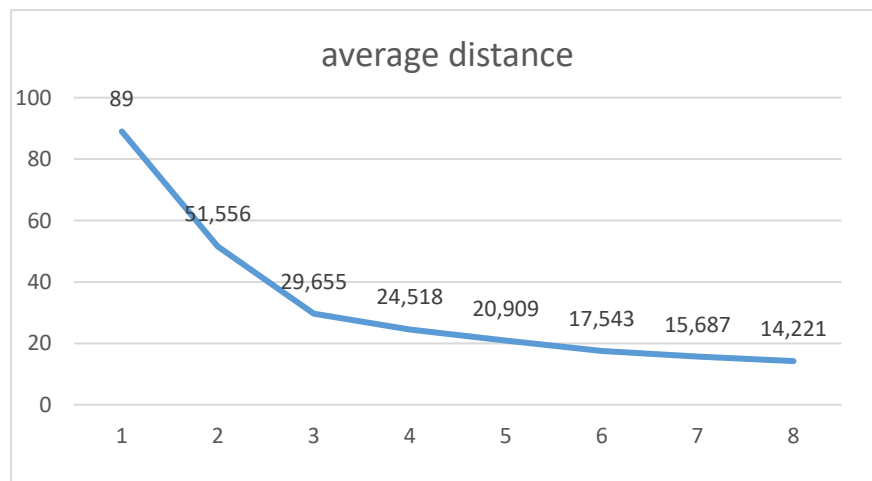
Tahap untuk menampilkan Hasil evaluasi data mining berdasarkan 3 cluster yang telah terbentuk.

4. Hasil dan Pembahasan

Studi ini memanfaatkan sebanyak 1275 entri data Penerimaan Mahasiswa Baru untuk tahun ajaran 2021/2022 dan 2022/2023, yang terdiri dari 5 atribut yaitu Jenis Kelamin, Asal Sekolah, Lokasi (Kecamatan), Lokasi (Kabupaten), dan Program Studi. Penelitian ini menggunakan perangkat lunak Rapidminer dengan format data berupa Comma Separated Values (.csv).

4.1 Uji Cluster Terbaik

Melakukan eksperimen untuk mencari nilai optimal k menggunakan Metode Elbow. Metode Elbow digunakan untuk menentukan jumlah kluster terbaik dengan memperhatikan persentase perbandingan antara jumlah kluster di mana titik siku akan terbentuk [3]. Hasil nilai optimal k kluster dari algoritma k-means dapat dilihat pada Gambar 3 seperti yang ditunjukkan di bawah ini.



Gambar 2. Hasil k optimal kluster algoritma k-means

Pada Gambar 3. Garis mengalami patahan yang membentuk siku (Elbow Method) pada $k = 3$. Artinya k optimal-nya adalah $k = 3$. Dari hasil optimasi diatas, peneliti menggunakan $k = 3$ yang akan digunakan dalam menganalisis pada Rapidminer..

4.2 Hasil Iterasi Metode K-Means

Adapun pengolahan data mahasiswa baru dengan menggunakan algoritma k-mean clustering dan software RepidMiner dapat dilihat pada gambar berikut:

Row No.	id	cluster	ixJenis Kel...	Prodi	Kecamatan	Kabupaten	Asal Sekolah
1	1	cluster_2	2	6	1	1	1
2	2	cluster_0	1	6	22	15	3
3	3	cluster_2	2	6	14	1	4
4	4	cluster_0	1	6	23	3	3
5	5	cluster_2	2	6	5	1	3
6	6	cluster_2	2	6	3	1	3
7	7	cluster_2	2	6	1	1	3
8	8	cluster_2	2	6	1	1	3
9	9	cluster_2	1	6	9	1	1
10	10	cluster_0	2	6	24	2	3
11	11	cluster_2	2	6	7	1	3
12	12	cluster_2	2	6	1	1	1
13	13	cluster_2	1	6	7	1	3
14	14	cluster_0	1	6	25	2	3
15	15	cluster_2	2	6	5	1	1
16	16	cluster_2	1	6	14	1	3
17	17	cluster_0	2	6	23	3	3
18	18	cluster_2	2	6	16	1	3

ExampleSet (1,275 examples, 2 special attributes, 5 regular attributes)

Gambar 3. Potongan Hasil Visualisasi Menggunakan Algoritma K-Means

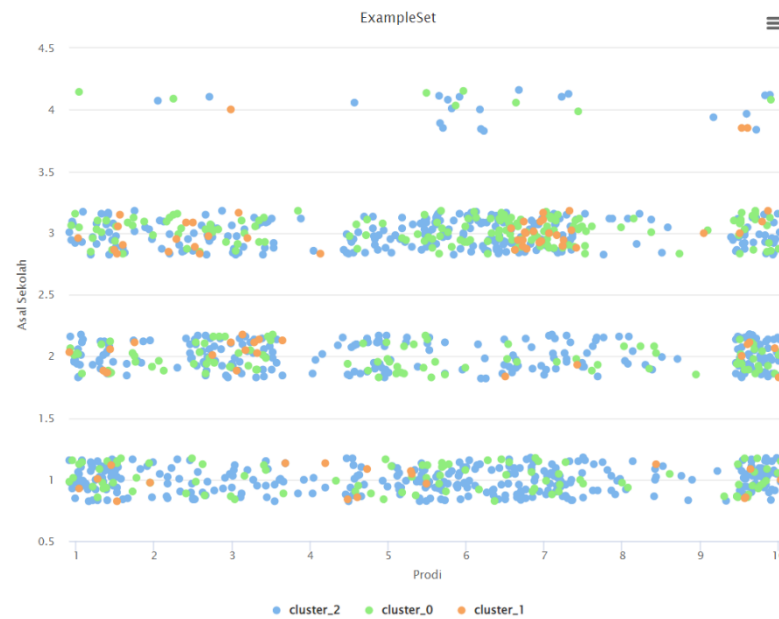
Dengan menggunakan pemodelan k-means clustering seperti gambar 2 diatas, dengan jumlah data 1275 dan inisialisasi jumlah cluster sebanyak 3 buah, sesuai dengan pendefinisian nilai k dengan jumlah cluster 0 : 354 items, cluster 1: 86 items, dan cluster 2 : 835 items.

Cluster Model

```
Cluster 0: 354 items
Cluster 1: 86 items
Cluster 2: 835 items
Total number of items: 1275
```

Gambar 4. Cluster Model

Hasil Penyebaran cluster 0, cluster 1, dan Cluster 2 sebanyak data 1275 pada pemodelan k means clustering dengan menggunakan repidminer, untuk 3 kelompok data dapat dilihat pada gambar berikut :



Gambar 5. Pemodelan K-Mean Clustering pada RapidMiner

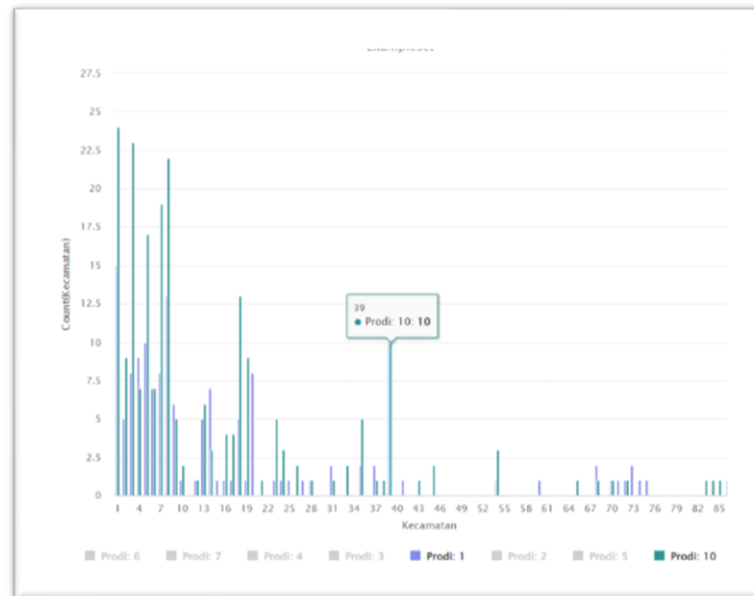
Visualisasi Persebaran ditandai 3 kelompok data. Kelompok pertama terlihat pada titik-titik penyebaran yang ditandai dengan warna hijau (cluster 0), kelompok kedua terlihat pada titik-titik penyebaran yang ditandai dengan warna orange (cluster 1), dan kelompok ketiga terlihat pada titik titik penyebaran yang ditandai dengan warna biru (cluster 2). Adapun hasil analisa cluster pada gambar 3, berisi tentang hasil pengelompokan berdasarkan kedekatan jarak antra titik pusat dengan data mahasiswa pada setiap atribut.

4.3 Evaluasi

berdasarkan hasil 3 Cluster diatas diperoleh hasil yang paling mendominasi adalah mahasiswa berjenis kelamin perempuan diikuti dengan laki – laki dengan asal Sekolah MA,SMA,SMK dengan prodi yang diambil adalah Pendidikan Guru Madrasah Ibtidaiyah,Informatika dan Pendidikan Agama Islam dengan distribusi wilayah Kec. Tertinggi adalah Kec.Buay Madang,Kec Belitang Madang Raya ,dan Kec.Belitang. Kabupaten tertinggi di wilayah OKU Timur, OKI dan OKU.

Sedangkan untuk mahasiswa terendah mengambil Program Studi Matematika, Fisika dan pertanian. Berdasarkan hasil diatas Kabupaten yang paling mendominasi adalah OKU Timur ,Kabupaten tersebut sama dengan wilayah Universitas Nurul Huda ,hal tersebut membuat keuntungan bagi team marketing untuk melakukan pemasaran di Kabupaten OKU Timur dengan distribusi 3 wilayah tertinggi di Kec.Buay Madang ,Kec.Belitang Madang Raya dan Kec.Belitang.

Hasil temuan berikutnya adalah terdapat wilayah potensial di luar Kabupaten OKU Timur yang memiliki basis pendaftar yang tinggi dengan 42 mahasiswa dengan paing banyak megambil prodi Informatika dan Pendidikan Bahasa Indonesia.Berikut grafik menampilkan wilayah potensial tersebut berada pada Ke.Mesuji Makmur Kab.OKI



Gambar 6. Grafik wilayah potensial Kec.Mesuji Makmur

4.4 Staretki Promosi

Berdasarkan hasil analisis, berikut adalah strategi promosi yang dapat diimplementasikan:

1. Penargetan Mahasiswa:
 - Memperkuat promosi untuk Program Studi Informatika dengan basis pendaftar tertinggi yaitu di SMK dan MA
 - Memperkuat promosi untuk Program Studi PGMI dan Pendidikan Agama Islam dengan basis pendaftar tertinggi yaitu di MA , Informatika, dan Pendidikan Agama Islam.
 - Mengoptimalkan promosi untuk tiga prodi dengan minat terendah yaitu Matematika,Pertanian dan Fisika
2. Penempatan Lokasi Promosi:
 - Mengutamakan wilayah dengan distribusi terbesar, seperti Kecamatan Buay Madang, Kecamatan Belitang Madang Raya, dan Kecamatan Belitang di Kabupaten OKU Timur.
 - Menyelenggarakan kegiatan promosi dan pameran di area-area yang menjadi basis pendaftar terbesar.
3. Peningkatan Visibilitas di Kabupaten OKU Timur:
 - Memanfaatkan keuntungan lokasi yang sama dengan wilayah Universitas Nurul Huda di Kabupaten OKU Timur.

- Mengoptimalkan strategi pemasaran di Kabupaten OKU Timur dengan fokus pada kegiatan pemasaran di Kecamatan Buay Madang, Kecamatan Belitang Madang Raya, dan Kecamatan Belitang.
- 4. Ekspansi ke Wilayah Potensial:
 - Menjangkau wilayah potensial di luar Kabupaten OKU Timur, seperti Kecamatan Mesuji Makmur di Kabupaten OKI dengan basis pendaftar tertinggi pada prodi Informatika dan Pendidikan Bahasa Indonesia.
 - Memperluas jangkauan promosi ke wilayah dengan basis pendaftar yang tinggi, terutama untuk prodi Informatika dan Pendidikan Guru Madrasah Ibtidaiyah dan Pendidikan Agama Islam.
- 5. Promotion Mix
 - Untuk wilayah yang jauh untuk dijangkau sesuai dengan temuan dari cluster 2 dan cluster 3, hanya dilakukan promosi melalui iklan yang tercantum dalam tabel promotion mix di bawah ini.

Tabel 13. Promotion Mix

No	Strategi Promosi	Cluster Satu	Cluster Dua	Cluster Tiga
1	Periklanan	√	√	√
2	Penjualan Personal	√	√	√
3	Promosi Penjualan	√		
4	Hubungan Masyarakat	√		
5	Pemasaran Langsung	√		

5. Kesimpulan

Hasil dari cluster data penerimaan mahasiswa baru tahun ajaran 2021/2022 dan 2022/2023 menunjukkan mahasiswa perempuan mendominasi diikuti oleh mahasiswa laki-laki, berasal dari Sekolah Menengah seperti MA, SMA, dan SMK, dengan minat pada program studi seperti Pendidikan Guru Madrasah Ibtidaiyah, Informatika, dan Pendidikan Agama Islam. Wilayah dengan distribusi tertinggi terdapat di Kecamatan Buay Madang, Kecamatan Belitang Madang Raya, dan Kecamatan Belitang di Kabupaten OKU Timur, OKI, dan OKU.

Sedangkan, mahasiswa dengan minat terendah mengambil Program Studi Matematika, Fisika, dan Pertanian. Kabupaten yang paling mendominasi adalah OKU Timur, yang juga merupakan wilayah dari Universitas Nurul Huda, memberikan keuntungan bagi tim pemasaran untuk fokus di wilayah tersebut, terutama di Kecamatan Buay Madang, Kecamatan Belitang Madang Raya, dan Kecamatan Belitang.

Ditemukan juga wilayah potensial di luar Kabupaten OKU Timur, khususnya di Ke. Mesuji Makmur, Kabupaten OKI, dengan jumlah pendaftar yang tinggi, terutama untuk program studi Informatika dan Pendidikan Bahasa Indonesia. serta untuk daerah yang jauh dijangkau sesuai dengan hasil cluster 2 dan cluster 3 hanya dilakukan pemasaran dengan periklanan yang tertera pada tabel *promotion mix*.

Acknowledgments

Ucapan terima kasih kepada tim peneliti dan Universitas Nurul Huda melalui Lembaga Penelitian dan Pengabdian kepada Masyarakat dalam pelaksanaan Hibah Internal tahun anggaran 2023 yang tertuang dalam Surat Kontrak No. 096/UNUHA/LPPM-PEN/XI/2023.

Daftar Pustaka

- [1] H. Hairani, D. Susilowati, I. Puji Lestari, K. Marzuki, and L. Z. A. Mardedi, "Segmentasi Lokasi Promosi Penerimaan Mahasiswa Baru Menggunakan Metode RFM dan K-Means Clustering," *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 21, no. 2, pp. 275–282, Mar. 2022, doi: 10.30812/matrik.v21i2.1542.
- [2] N. Afriani Manihuruk, M. Zarlis, E. Irawan, and H. Satria Tambunan, "Penerapan Data Mining Dalam Mengelompokkan Calon Penerima Beasiswa Dengan Menggunakan Algoritma K-Means," vol. 4, no. 1, 2020, doi: 10.30865/komik.v4i1.2575.
- [3] F. Nasari and S. Darma, "PENERAPAN K-MEANS CLUSTERING PADA DATA PENERIMAAN MAHASISWA BARU (STUDI KASUS : UNIVERSITAS POTENSI UTAMA)," pp. 6–8, 2015.
- [4] Johan Oscar Ong, "IMPLEMENTASI ALGORITMA K-MEANS CLUSTERING UNTUK MENENTUKAN STRATEGI MARKETING PRESIDENT UNIVERSITY," vol. 12, no. 1, Jun. 2013.
- [5] D. Exasanti and A. Jananto, "Analisa Hasil Pengelompokan Wilayah Kejadian Non-Kebakaran Menggunakan Agglomerative Hierarchical Clustering pada Dinas Pemadam Kebakaran Kota Semarang."
- [6] W. Purba and W. Siawin, "IMPLEMENTASI DATA MINING UNTUK PENGELOMPOKKAN DAN PREDIKSI KARYAWAN YANG BERPOTENSI PHK DENGAN ALGORITMA K-MEANS CLUSTERING," *Jurnal Sistem Informasi Ilmu Komputer Prima (JUSIKOM PRIMA)*, vol. 2, no. 2, 2019.
- [7] F. A. Tanjung, A. P. Windarto, M. Fauzan, M. P. Studi, S. Informasi, and S. Tunas Bangsa, "Penerapan Metode K-Means Pada Pengelompokan Pengangguran Di Indonesia," vol. 6, pp. 61–74, [Online]. Available: <https://tunasbangsa.ac.id/ejurnal/index.php/jurasik>
- [8] Y. P. Pasrun, "IMPLEMETASI K-MEANS CLUSTERING PADA RAPIDMINER UNTUK ANALISIS DAERAH RAWAN KECELAKAAN," 2017. [Online]. Available: <https://www.researchgate.net/publication/316661764>
- [9] N. Lelawati, S. Retnaning Rahayu, and U. Muhammadiyah Metro, "STRATEGI PROMOSI PENERIMAAN MAHASISWA BARU (PENMARU) UM METRO," *Jurnal Lentera Pendidikan Pusat Penelitian LPPM UM METRO*, vol. 6, no. 2, 2021.
- [10] R. Wulan Sari, A. Wanto, and A. Perdana Windarto, "KOMIK (Konferensi Nasional

Teknologi Informasi dan Komputer) IMPLEMENTASI RAPIDMINER DENGAN METODE K-MEANS (STUDY KASUS: IMUNISASI CAMPAK PADA BALITA BERDASARKAN PROVINSI)", [Online]. Available: <http://ejurnal.stmik-budidarma.ac.id/index.php/komik>

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>)