

Penerapan Metode K-Means Clustering Dalam Menentukan Strategi Promosi Berdasarkan Data Penerimaan Mahasiswa Baru

Nanda Ayu Rahmalinda¹, Arief Jananto²

^{1,2} Fakultas Teknologi Informasi dan Industri, Sistem Informasi, Universitas Stikubank, Semarang, Indonesia
Jl. Tri Lomba Juang No 1 Semarang 50241, Kota Semarang, Jawa Tengah
Email: nandaayu762@gmail.com, ajananto09@edu.unisbank.ac.id

Abstrak – Banyaknya perguruan tinggi yang tersebar di berbagai daerah di tanah air, seharusnya mampu untuk membentuk karakteristik manusia dan meningkatkan kualitas pendidikan di Indonesia. Sekolah Tinggi Ilmu Ekonomi Assholeh Pemalang merupakan suatu institusi atau lembaga pendidikan yang berbentuk Perguruan Tinggi Swasta (PTS) dengan tiga program studi yaitu S1-Manajemen, S1-Akuntansi dan D3-Manajemen Perusahaan, yang beralamat di Kabupaten Pemalang, Jawa Tengah. Data yang digunakan dalam penelitian ini merupakan data dari hasil proses Penerimaan Mahasiswa Baru yang selanjutnya akan dilakukan proses pengolahan dan menghasilkan suatu hasil analisa menggunakan *data mining*. Dengan tujuan untuk menentukan strategi promosi berdasarkan daerah asal dari profil mahasiswa baru yang mendaftar. Metode yang digunakan adalah *Clustering* dan algoritma yang digunakan adalah *K-Means*. Data yang sudah diolah akan dilakukan proses analisa menggunakan teknik *Knowledge Discovery in Databases* (KDD) dengan lima tahapan yaitu *selection*, *pre-processing*, *transformation*, *data mining* dan *evaluation*. Proses implementasi dalam penelitian ini menggunakan *software* WEKA versi 3.8.5, dan menghasilkan tiga klasterisasi yaitu *Cluster 1* berjumlah 189 data dengan presentase 42%, *Cluster 2* berjumlah 186 data dengan presentase 41% dan *Cluster 3* berjumlah 78 data dengan presentase 17%. Dengan kecepatan komputasinya sebesar 0.01 detik. Proses promosi dilakukan dengan mengirimkan *team marketing* di persebaran wilayah pada kabupaten yang mendominasi dengan cara mendatangi langsung dan melakukan sosialisasi untuk mengenalkan Sekolah Tinggi Ilmu Ekonomi Assholeh Pemalang dikalangan siswa/i, membagikan brosur dan menempelkan pamflet serta melakukan penyalarsan menggunakan strategi *promotion mix*.

Kata Kunci: Data Mining, *K-Means Clustering*, Penerimaan Mahasiswa Baru, Strategi Promosi, WEKA

Abstract - The number of universities spread across various regions in the country should be able to shape human characteristics and improve the quality of education in Indonesia. Assholeh School of Economics Pemalang is an institution or educational institution in the form of a Private Higher Education (PTS) with three study programs, namely S1-Management, S1-Accounting and D3-Company Management, which is located in Pemalang Regency, Central Java. The data used in this study is data from the results of the New Student Admission process which will then be processed and produce an analysis using data mining. With the aim of determining the promotion strategy based on the area of origin of the profile of new students who register. The method used is Clustering and the algorithm used is K-Means. The processed data will be analyzed using the Knowledge Discovery in Databases (KDD) technique with five stages, namely selection, pre-processing, transformation, data mining and evaluation. The implementation process in this study uses WEKA software version 3.8.5, and produces three clusters, namely Cluster 1 totaling 189 data with a percentage of 42%, Cluster 2 amounting to 186 data with a percentage of 41%, and Cluster 3 totaling 78 data with a percentage of 17%. With a computing speed of 0.01 seconds. The promotion process is carried out by sending a marketing team in the distribution of areas in the dominating districts by visiting directly and conducting socialization to introduce the Assholeh Pemalang High School of Economics among students, distributing brochures and pasting pamphlets and aligning using the promotion mix strategy.

Keywords: Data Mining, K-Means Clustering, New Student Admission, Promotion Strategy, WEKA

1. PENDAHULUAN

1.1 Latar Belakang

Dalam dunia pendidikan, khususnya di perguruan tinggi baik perguruan tinggi negeri maupun swasta, proses penerimaan mahasiswa baru merupakan hal yang sangat penting dan wajib dilakukan di setiap tahunnya. Dimana pendidikan sendiri merupakan proses untuk membentuk karakteristik dan menciptakan mutu yang berkualitas [1]. sedangkan perguruan tinggi merupakan tempat yang digunakan untuk meningkatkan mutu dengan cara sederhana namun penting dan standar yang harus sesuai dengan tujuannya [2]. Sekolah Tinggi Ilmu Ekonomi Assholeh

Pemalang merupakan suatu institusi Pendidikan Tinggi yang berbentuk Perguruan Tinggi Swasta dengan tiga program studi yaitu S1-Manajemen, S1-Akuntansi dan D3-Manajemen Perusahaan yang beralamat di Jalan Sindoro no.39, Mulyoharjo, Kabupaten Pemalang, Jawa Tengah. Dalam proses Penerimaan Mahasiswa Baru di Sekolah Tinggi Ilmu Ekonomi Assholeh Pemalang tentunya memerlukan strategi promosi dengan tujuan untuk mendapatkan banyak calon mahasiswa baru baik mahasiswa dari dalam maupun luar kabupaten/kota. Proses Penerimaan Mahasiswa Baru menghasilkan banyak data berupa profil dari mahasiswa yang mendaftar. Dimana dalam proses Penerimaan Mahasiswa Baru tersebut, pihak Sekolah Tinggi Ilmu Ekonomi Assholeh Pemalang, sudah melakukan promosi dengan cara memasang baliho, selebaran, memberikan potongan biaya dan kunjungan ke SMA/SMK di Kabupaten Pemalang. Jika data yang diperoleh di olah dengan proses yang tepat, maka informasi tersembunyi dari profil data mahasiswa dapat diketahui dan bermanfaat bagi pihak perguruan tinggi untuk memudahkan dalam menentukan strategi promosi baru berdasarkan pengelompokan daerah asal mahasiswa.

Dalam penelitian sebelumnya yang dilakukan oleh Manihuruk dkk untuk membantu pengelompokan calon penerima beasiswa menggunakan 3 cluster dengan hasil penelitian diperoleh cluster 1 sebanyak 73 orang, cluster 2 sebanyak 30 orang dan cluster 3 sebanyak 25 orang[3]. Nasari dan Darma untuk menentukan strategi promosi berdasarkan asal sekolah dan jurusan yang diambil menggunakan 2 klaster dan diperoleh hasil pada cluster 1 jika asal sekolah "SMA" maka rata-rata jurusan yang diambil "Sistem Informasi" dan pada cluster 2 jika asal sekolah "SMK" maka rata-rata jurusan yang diambil "Teknik Informatika"[4]. Ong melakukan penelitian menggunakan 3 klaster untuk membantu pihak marketing dalam melakukan promosi berdasarkan data mahasiswa yang telah lulus tahun 2012 dan diperoleh hasil cluster 1 mahasiswa yang berasal dari DKI Jakarta dan Jawa Barat mengambil jurusan *Information Technology* dan *Marketing*, cluster 2 mahasiswa berasal dari DKI Jakarta dan Jawa Barat mengambil jurusan *Accounting* dan *International Business* dan cluster 3 mahasiswa yang berasal dari Sulawesi, Jawa Timur dan Sumatera mengambil jurusan *Public Relation*, *Accounting* dan *International Business*[5]. Arifin melakukan penelitian menggunakan 2 klaster untuk membantu pihak sekolah dalam mengarahkan siswa/siswi dalam memilih Perguruan Tinggi Negeri yang di minati dan diperoleh hasil pada cluster 1 sebanyak 46 siswa dengan UNPAD = 23 siswa, ITB = 20 siswa, UNDIP = 2 siswa dan UPI = 1 siswa, pada cluster 2 sebanyak 40 siswa dengan UNPAD = 24 siswa, UPI = 1 siswa, UGM = 1 siswa, UNSRI = 1 siswa dan ISI = 1 siswa [6]. Peneliti tertarik untuk mengelompokkan daerah asal mahasiswa baru di Sekolah Tinggi Ilmu Ekonomi Assholeh Pemalang menggunakan metode *K-Means Clustering* dengan *software* yang digunakan yaitu WEKA (*Waikato Environment for Knowledge Analysis*) untuk mempermudah dalam memberikan gambaran yang jelas dari hasil penelitian yang dilakukan. Maka judul dari penelitian ini adalah Penerapan Metode *K-Means Clustering* Dalam Menentukan Strategi Promosi Berdasarkan Data Penerimaan Mahasiswa Baru.

Perumusan masalah dalam penelitian ini adalah bagaimana menerapkan metode *K-Means Clustering* untuk menentukan strategi promosi berdasarkan data penerimaan mahasiswa baru di Sekolah Tinggi Ilmu Ekonomi Assholeh Pemalang.

1.2 Tujuan dan Manfaat Penelitian

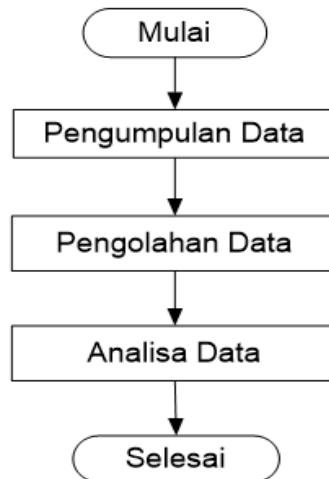
Tujuan yang ingin dicapai dalam penelitian ini adalah untuk menghasilkan pengelompokan data mahasiswa baru menggunakan teknik data mining berdasarkan profil mahasiswa baru yang mendaftar di Sekolah Tinggi Ilmu Ekonomi Assholeh Pemalang.

Manfaat yang diharapkan dalam penelitian ini adalah untuk mengetahui pola daerah hasil mahasiswa baru sehingga didapatkan informasi dan memudahkan pihak Sekolah Tinggi Ilmu Ekonomi Assholeh Pemalang dalam menentukan strategi promosi secara tepat.

2. METODE PENELITIAN

2.1 Tahapan Penelitian

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



Gambar 1. Tahapan Penelitian

2.1.1 Tahap Pengumpulan Data

Dalam tahap ini peneliti melakukan pengumpulan data di Sekolah Tinggi Ilmu Ekonomi Assholeh Pemalang. Data yang dikumpulkan berupa data dari proses Penerimaan Mahasiswa Baru dengan Tahun Ajaran 2018/201, 2019/2020, 2020/2021, dan 2021/2022.

2.1.2 Tahap Pengolahan Data

Data yang dikumpulkan masih berupa data mentah yang selanjutnya akan dilakukan proses editing, seperti menghapus data yang kosong, melengkapi data yang kurang jelas dan kesalahan yang ada pada data.

2.1.3 Tahap Analisa Data

Metode yang digunakan untuk proses analisa adalah *Knowledge Discovery in Databases* (KDD). Dimana metode tersebut merupakan proses *non-trivial* untuk menemukan pola yang bersifat baru, berguna dan mudah dipahami[7]. dengan lima proses yang berurutan yaitu *selection*, *pre-processing*, *transformation*, *data mining*, dan *evaluation* [8].

a. *Selection* (Seleksi Data)

Proses ini data di ambil dan diseleksi yang kemudian dilakukan proses analisa agar dapat dilakukan pemrosesan data minig. Hasil seleksi data dapat dilihat pada Tabel 1.

Tabel 1. Potongan Hasil Seleksi Data

No	Jenis Kelamin	Asal Sekolah	Alamat Mahasiswa	Program Studi
1	P	SMK Texmaco Pemalang	Jl Markisa RT 001 RW 005 Tambakrejo Kec. Pemalang, Kab. Pemalang	S1-Manajemen
2	L	PKBM Agung Nugraha	Gg Kepodang RT 06 RW 02 Desa Kramat Kec. Pemalang Kab. Pemalang	S1-Manajemen
3	P	SMK Amanah Husada	Demangharjo RT 03 RW 04 Kec. Warureja, Kab.Tegal	S1-Manajemen
4	P	SMA N 2 Pemalang	Jl. Kulbanda Gg Kenanga RT 01 RW 08 Beji babadan, Kec. Taman, Kab. Pemalang	S1-Manajemen
5	L	MAN Pemalang	Jl Cempaka Ds Wanarejan utara RT 01 RW 02 Kec. Taman Pemalang	S1-Manajemen

Dari Tabel 1 di atas, dipilih empat atribut yang akan digunakan. Karena dalam penelitian ini akan menentukan strategi promosi PMB berdasarkan banyaknya daerah asal (Kabupaten/Kota) dari mahasiswa baru yang mendaftar di Sekolah Tinggi Ilmu Ekonomi Assholeh Pemalang pada Tahun Ajaran 2018/2019, 2019/2020, 2020/2021, dan 2021/2022.

b. *Pre-processing*

Proses untuk menghilangkan *noise*, data duplikasi data yang *in-konsisten* dan data yang kosong dari data yang ada.

c. *Transformation* (Transformasi Data)

Proses untuk melakukan transformasi dan konversi terhadap data yang belum mempunyai entitas yang jelas.

1. Transformasi

Atribut yang akan ditransformasikan sebagai berikut:

1.) Atribut Asal Sekolah

Data pada atribut Asal Sekolah berupa identitas sekolah sebelumnya. Dari banyaknya asal sekolah tersebut, Adapun dikelompokkan menjadi 4 kategori : “SMA”, “SMK”, “MA” dan “Lainnya”. Untuk lebih jelasnya dapat dilihat pada Tabel 2 dibawah ini.

Tabel 2. Transformasi Atribut Asal Sekolah

No.	Asal Sekolah	Hasil Transformasi
1	SMA N 2 Pemalang	SMA
2	SMK Texmaco Pemalang	SMK
3	SMK Amanah Husada	
4	MAN Pemalang	MA
5	PKBM Agung Nugraha	Lainnya

2.) Atribut Alamat Mahasiswa

Data pada atribut Alamat Mahasiswa berupa daerah berdasarkan kecamatan dari masing-masing mahasiswa. Dari beberapa kecamatan tersebut, adapun dikelompokkan menjadi 3 kabupaten : Kabupaten Pemalang, Kabupaten Tegal dan Kabupaten Brebes. Untuk lebih jelasnya dapat dilihat pada Tabel 3 dibawah ini.

Tabel 3. Transformasi Atribut Alamat Mahasiswa

No.	Alamat Mahasiswa (Kecamatan)	Hasil Transformasi
1	Kec. Ampelgading	
2	Kec. Bantar Bolang	
3	Kec. Belik	
4	Kec. Bodeh	
5	Kec. Comal	
6	Kec. Moga	
7	Kec. Pemalang	Pemalang
8	Kec. Petarukan	
9	Kec. Pulosari	
10	Kec. Randudongkal	
11	Kec. Taman	
12	Kec. Ulujami	
13	Kec. Warungpring	
14	Kec. Watukumpul	
15	Kec. Warureja	
16	Kec. Kedungbanteng	
17	Kec. Kramat	Tegal
18	Kec. Adiwerna	
19	Kec. Suradadi	
20	Kec. Bumiayu	
21	Kec. Jatibarang	Brebes
22	Kec. Brebes	

2. Konversi

Data yang sudah di transformasi kemudian dilakukan proses konversi dengan tujuan untuk mempermudah dalam proses data mining. Atribut yang akan di konversi sebagai berikut:

1.) Jenis Kelamin

Data pada atribut Jenis Kelamin adalah L dan P. Hasil konversi dapat dilihat pada Tabel 4.

Tabel 4. Konversi Atribut Jenis Kelamin

Jenis Kelamin	Hasil Konversi
L	1
P	2

2.) Asal Sekolah

Data pada atribut Asal Sekolah adalah SMA, SMK, MA dan Lainnya. Hasil konversi dapat dilihat pada Tabel 5.

Tabel 5. Konversi Atribut Asal Sekolah

Asal Sekolah	Hasil Konversi
SMA	1
SMK	2
MA	3
Lainnya	4

3.) Alamat Mahasiswa

Data pada atribut Alamat Mahasiswa adalah Kab. Pemasang, Kab. Tegal dan Kab. Brebes. Hasil konversi dapat dilihat pada Tabel 6.

Tabel 6. Konversi Atribut Alamat Mahasiswa

Alamat Mahasiswa (Kabupaten)	Hasil Konversi
Pemasang	1
Tegal	2
Brebes	3

4.) Program Studi

Data pada atribut Program Studi adalah S1-Manajemen, S1-Akuntansi dan D3-Manajemen Perusahaan. Hasil konversi dapat dilihat pada Tabel 7.

Tabel 7. Konversi Atribut Program Studi

Program Studi	Hasil Konversi
S1-Manajemen	1
S1-Akuntansi	2
D3-Manajemen Perusahaan	3

Hasil data dari keseluruhan proses *pre-processing* dapat dilihat pada Tabel 8.

Tabel 8. Potongan Data Setelah Proses Konversi

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

d. *Data Mining*

Data Mining adalah proses menentukan pola atau informasi menarik dari sekumpulan data dengan teknik, metode atau algoritma tertentu[9] dan menghasilkan pengetahuan baru yang berasal dari data lama yang berguna dalam pengambilan dan menentukan keputusan di masa yang akan datang[10]. Tahap untuk membentuk suatu pola dari data yang ada menggunakan metode *clustering*. *Clustering* adalah algoritma pengelompokan data dengan jumlah besar di proses menjadi bagian kecil dengan atribut kemiripan tertentu[8]. Analisis *clustering* dilakukan dengan membagi data dalam satu set menjadi beberapa bagian dengan kesamaan dalam satu kelompok lebih besar dari data yang lain[11]. Dalam penelitian ini menggunakan algoritma *K-Means Clustering*. *K-Means clustering* adalah salah satu metode dari proses data mining dimana mengelompokkan data dengan karakteristik yang sama antara satu sama lain dan memiliki karakteristik yang berbeda dengan data dari kelompok lain[4]. Dasar perhitungan dari algoritma *K-means* adalah sebagai berikut:

1. Menentukan nilai k yang digunakan sebagai jumlah kluster yang ingin dibentuk.
 Dalam penelitian ini jumlah *centroid* yang digunakan sebanyak 3 *cluster*.
2. Menginisialisasi k pusat kluster atau *centroid* yang dapat dilakukan secara acak.
 Titik *cluster* yang telah dipilih secara acak dapat dilihat pada Tabel 9.

Tabel 9. Titik Pusat *Cluster*

Atribut	Cluster 1	Cluster 2	Cluster 3
Jenis_Kelamin	2	1	2
Asal_Sekolah	1	2	2
Alamat_Mahasiswa	1	2	1
Program_Studi	1	1	2

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

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

Keterangan:
 d = titik dokumen
 P = data record
 Q = data *centroid*

Contoh perhitungan Data ke-1 dapat dilihat sebagai berikut:

$$d(1,1) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 + (x_4 - y_4)^2}$$

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

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

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

$$d(1,2) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 + (x_4 - y_4)^2}$$

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

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

$$d(1,2) = 1,4123562$$

$$d(1,3) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 + (x_4 - y_4)^2}$$

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

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

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

4. Mengelompokkan setiap data berdasarkan jarak terdekat antara data dengan *centroid*-nya.
 Hasil dari perhitungan jarak menggunakan rumus *Eucliden Distance* dapat dilihat pada Tabel 10.

Tabel 10. Hasil Perhitungan Jarak Terdekat

Data ke-	Jenis Kelamin	Asal Sekolah	Alamat Mahasiswa	Program Studi	Jarak ke K1	Jarak ke K2	Jarak ke K3	Cluster
1	2	2	1	1	1	1,414213562	1	Cluster 1
2	1	4	1	1	3,16227766	2,236067977	2,449489743	Cluster 2
3	2	2	1	1	1	1,414213562	1	Cluster 1
4	2	2	1	1	1	1,414213562	1	Cluster 1
5	2	2	1	1	1	1,414213562	1	Cluster 1

5. Tentukan posisi *centroid* baru (k). Lakukan kembali langkah nomer 3 jika pusat klaster masih berubah sampai pusat klaster tidak berubah lagi.
 Hasil Akhir Perhitungan Cluster dapat dilihat pada Tabel 11.

Tabel 11. Hasil Akhir Perhitungan Cluster

Data ke-	Jenis Kelamin	Asal Sekolah	Alamat Mahasiswa	Program Studi	Hasil Cluster Iterasi ke-1	Hasil Cluster Iterasi ke-2	Hasil Cluster Iterasi ke-3	Hasil Cluster Iterasi ke-4
1	2	2	1	1	Cluster 1	Cluster 1	Cluster 1	Cluster 1
2	1	4	1	1	Cluster 2	Cluster 2	Cluster 2	Cluster 2
3	2	2	1	1	Cluster 1	Cluster 1	Cluster 1	Cluster 1
4	2	2	1	1	Cluster 1	Cluster 1	Cluster 1	Cluster 1
5	2	2	1	1	Cluster 1	Cluster 1	Cluster 1	Cluster 1

- e. Interpretasi/Evaluasi
 Tahap untuk menampilkan *output* dari pemrosesan *Data Mining*.

2.2 Landasan Teori

a. Software WEKA

Waikato Environment for Knowledge Analysis (WEKA) adalah aplikasi *open source* yang berbasis *java* dan dikembangkan oleh Universitas Waikato di Selandia Baru, berisikan kumpulan algoritma serta visualisasi untuk analisis data dan pemodelan[9]. Dalam penelitian ini menggunakan *software* WEKA versi 3.8.5. dapat dilihat pada Gambar 2 sebagai berikut.

b. Promosi

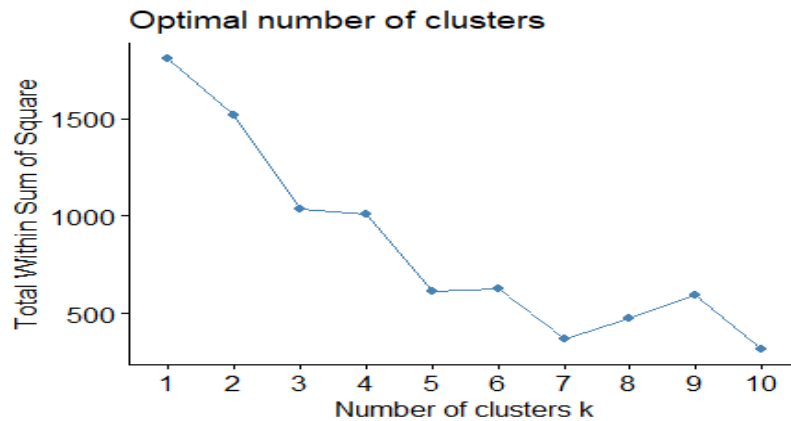
Promosi adalah kegiatan dalam membangun suatu komunikasi antara pihak intitusi dengan calon mahasiswa baru dan keluarganya dengan tujuan untuk mempengaruhi keputusan dalam memilih suatu institusi Pendidikan[12]. Ada 5 variabel *promotion mix* yaitu : periklanan, penjualan personal, promosi penjualan, hubungan masyarakat dan pemasaran langsung. Bauran promosi adalah kombinasi variabel dalam *promotion mix* dengan tujuan untuk mencapai target dari program penjualan yang sudah direncanakan[13].

3. HASIL DAN PEMBAHASAN

Penelitian ini menggunakan jumlah data sebanyak 453 *record* data Penerimaan Mahasiswa Baru dengan empat atribut yaitu Jenis Kelamin, Asal Sekolah, Alamat Mahasiswa dan Program Studi. Dalam penelitian ini menggunakan *software* WEKA dengan format data *Attribute-Relation File Format* (.arff) atau *Comma Separated Values* (.csv).

3.1 Uji Cluster Terbaik

Melakukan uji coba untuk mencari *k* optimal menggunakan *Elbow Method*. *Elbow Method* adalah suatu metode yang dipakai untuk menentukan jumlah *cluster* terbaik dengan melihat persentase perbandingan antara jumlah cluster dimana akan membentuk siku pada suatu titik [14]. Hasil *k* optimal klaster algoritma *k-means* dapat dilihat pada Gambar 3 sebagai berikut.



Gambar 3. Hasil k optimal kluster algoritma k -means

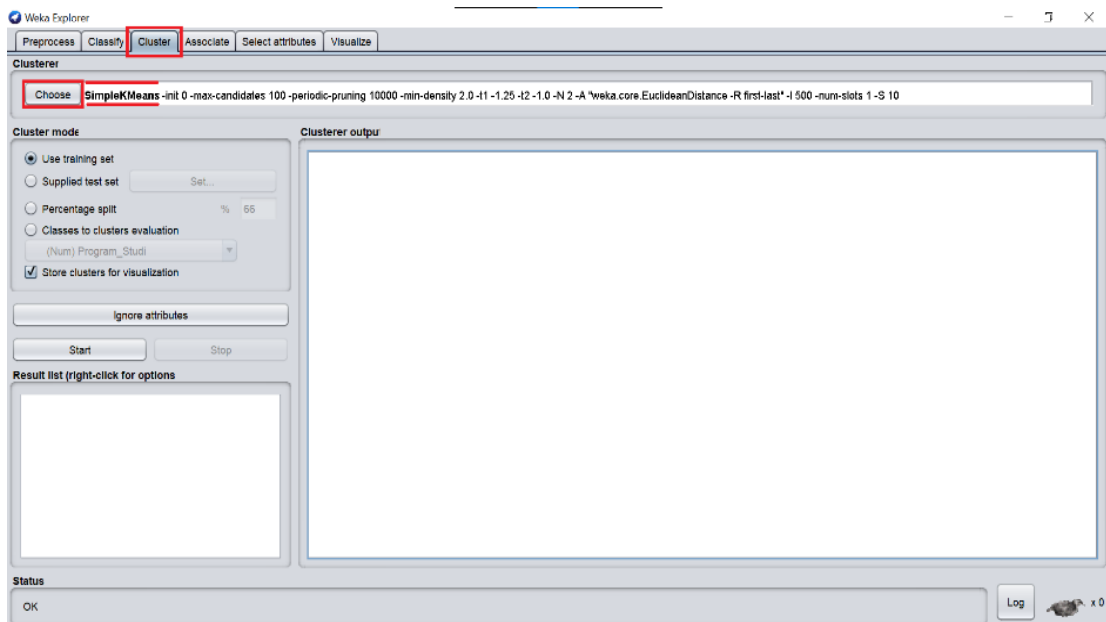
Pada Gambar 3. Garis mengalami patahan yang membentuk siku (*Elbow Method*) pada $k = 3$. Artinya k optimal-nya adalah $k = 3$. Dari hasil optimasi diatas, peneliti menggunakan $k = 3$ yang akan digunakan dalam menganalisis pada WEKA.

3.2 Implementasi software WEKA

Data yang di-cluster-kan sebanyak 453 *record* yang akan dibentuk menjadi 3 *cluster* dengan metode perhitungan jarak dari setiap data ke pusat *cluster* menggunakan persamaan rumus *Euclidean Distance* dan algoritma *K-Means Clustering*.

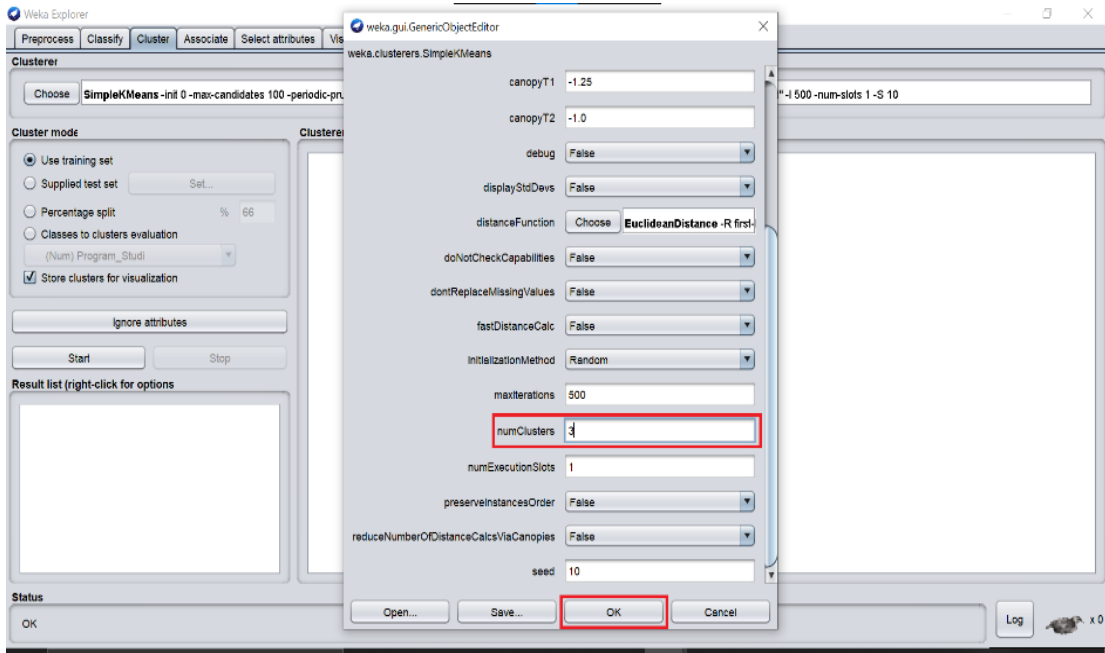
Software WEKA yang digunakan adalah WEKA versi 3.8.5. Untuk proses pengimplementasian dapat dilihat sebagai berikut.

1. Klik “cluster” lalu klik “choose” untuk memilih algoritma yang dipakai. Dalam pengimplementasian software WEKA disini menggunakan algoritma “SimpleKmeans”. Dapat dilihat pada Gambar 4. Sebagai berikut.



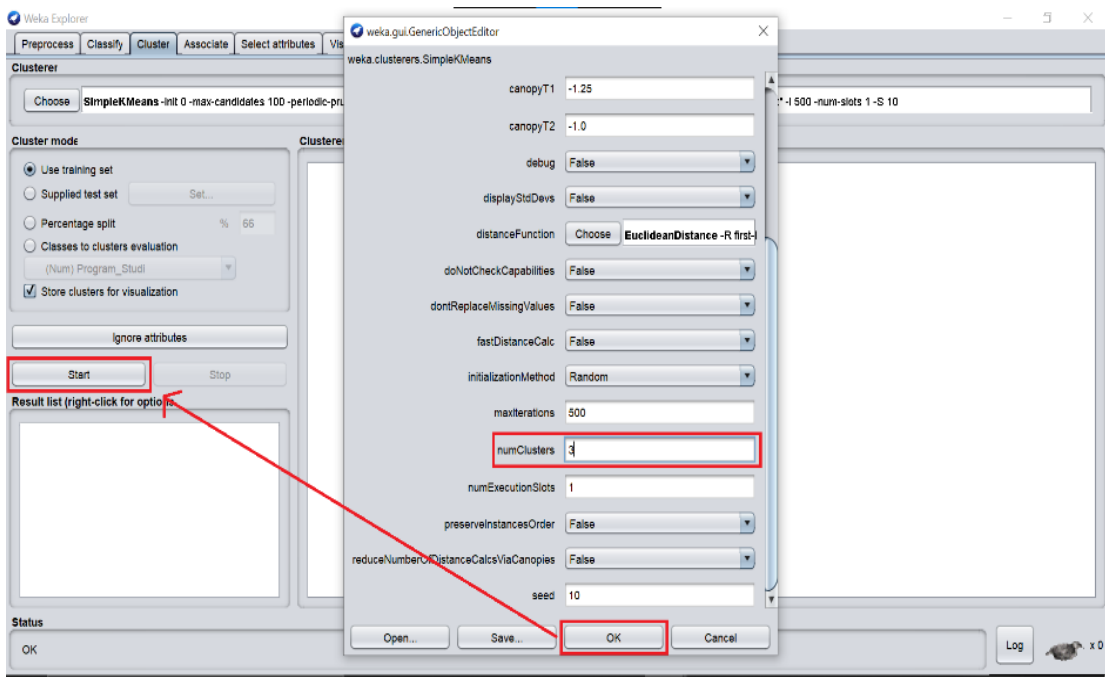
Gambar 4. Tampilan Pada Metode Cluster

2. *Double Click* pada bagian yang ditandai dan akan muncul tampilan parameter pengolahan data. Pastikan pada *distanceFunction* pilih “EuclideanDistance -R first-last” dan kolom *numClusters* diisi sesuai dengan jumlah *cluster* yang akan digunakan, lalu klik *OK*. Dapat dilihat pada Gambar 5 sebagai berikut.

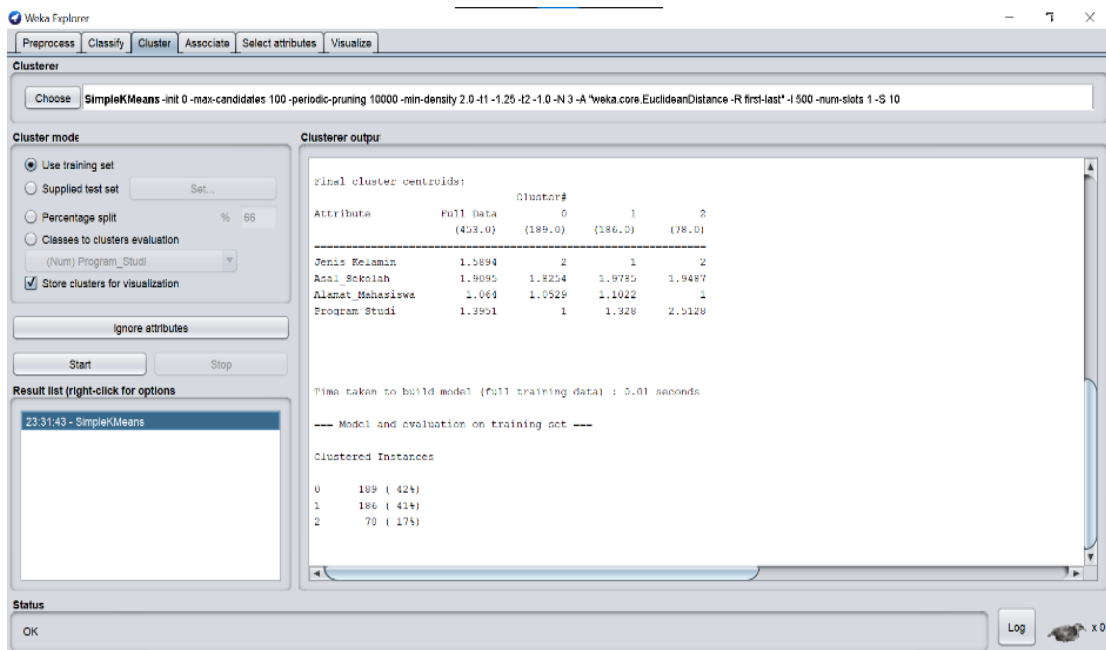


Gambar 5. Tampilan Editor *SimpleKmeans*

3. Klik *Start* (pada Gambar 6) untuk melakukan pengolahan data dan menghasilkan *output*, seperti pada tampilan *Clusterer Output* (Gambar 7) sebagai berikut.



Gambar 6. Tampilan Pengolahan Data.



Gambar 6. Hasil Clustering Menggunakan Software WEKA

Tabel 9. Hasil Clustering Dengan Tiga Cluster

Cluster Ke- (dalam software WEKA)	Jumlah Anggota Cluster
Cluster 0	189
Cluster 1	186
Cluster 2	78

Keterangan :

- a. Cluster 0 yaitu Cluster 1
- b. Cluster 1 yaitu Cluster 2
- c. Cluster 2 yaitu Cluster 3

Jumlah keseluruhan data yang telah di uji menggunakan *software WEKA* sebanyak 453 *record* dengan 3 *cluster* di peroleh yaitu : *Cluster 1* berjumlah 189 (42%), *Cluster 2* berjumlah 186 (41%) dan *Cluster 3* berjumlah 78 (17%). Dengan kecepatan komputasi sebesar 0.01 detik.

Data yang telah di uji dengan menggunakan algoritma *K-Means Clustering* mendapatkan hasil *plot*, dapat dilihat pada Gambar 6 sebagai berikut.



Gambar 6. Hasil Plot Menggunakan Algoritma K-Means Clustering

The figure shows a screenshot of the ARFF Viewer displaying the results of the K-Means clustering. The table contains 35 rows of data, each representing an instance. The columns are: Instance_number (1-35), Jenis_Kelamin (0.0 or 1.0), Asal_Sekolah (2.0 or 3.0), Alamat_Mahasiswa (2.0 or 4.0), Program_Studi (1.0 or 2.0), and Cluster (cluster0 or cluster1).

No	1: Instance_number	2: Jenis_Kelamin	3: Asal_Sekolah	4: Alamat_Mahasiswa	5: Program_Studi	6: Cluster
	Numeric	Numeric	Numeric	Numeric	Numeric	Nominal
1	0.0	2.0	2.0	1.0	1.0	cluster0
2	1.0	1.0	4.0	1.0	1.0	cluster1
3	2.0	2.0	2.0	1.0	1.0	cluster0
4	3.0	2.0	2.0	1.0	1.0	cluster0
5	4.0	2.0	2.0	1.0	1.0	cluster0
6	5.0	1.0	2.0	1.0	1.0	cluster1
7	6.0	2.0	2.0	1.0	1.0	cluster0
8	7.0	2.0	3.0	1.0	1.0	cluster0
9	8.0	2.0	1.0	1.0	1.0	cluster0
10	9.0	1.0	2.0	1.0	1.0	cluster1
11	10.0	1.0	4.0	1.0	1.0	cluster1
12	11.0	2.0	3.0	1.0	1.0	cluster0
13	12.0	2.0	4.0	1.0	1.0	cluster0
14	13.0	1.0	2.0	1.0	1.0	cluster1
15	14.0	2.0	2.0	1.0	1.0	cluster0
16	15.0	1.0	2.0	1.0	1.0	cluster1
17	16.0	2.0	2.0	1.0	1.0	cluster0
18	17.0	1.0	2.0	1.0	1.0	cluster1
19	18.0	1.0	2.0	1.0	1.0	cluster1
20	19.0	2.0	1.0	1.0	1.0	cluster0
21	20.0	1.0	4.0	2.0	1.0	cluster1
22	21.0	1.0	2.0	1.0	1.0	cluster1
23	22.0	2.0	1.0	1.0	1.0	cluster0
24	23.0	2.0	2.0	1.0	1.0	cluster0
25	24.0	2.0	2.0	1.0	1.0	cluster0
26	25.0	1.0	2.0	1.0	1.0	cluster0
27	26.0	2.0	1.0	1.0	1.0	cluster0
28	27.0	1.0	1.0	1.0	1.0	cluster1
29	28.0	1.0	2.0	1.0	1.0	cluster1
30	29.0	2.0	1.0	1.0	1.0	cluster0
31	30.0	2.0	1.0	1.0	1.0	cluster0
32	31.0	1.0	1.0	1.0	1.0	cluster1
33	32.0	2.0	2.0	1.0	1.0	cluster0
34	33.0	2.0	1.0	1.0	1.0	cluster0
35	34.0	2.0	1.0	1.0	1.0	cluster0

Gambar 7. Potongan Hasil Visualisasi Menggunakan Algoritma K-Means Clustering

3.3 Hasil Analisa Uji Coba Software WEKA

Hasil analisa dari uji coba menggunakan WEKA 3.8.5 dengan jumlah data sebanyak 453 *record* dan 3 *cluster*, didapatkan jumlah anggota *Cluster 1* sebanyak 189 data. Pada atribut “Jenis Kelamin” yang mendominasi adalah P (Perempuan) sebanyak 189 orang. Pada atribut “Asal Sekolah” yang mendominasi adalah SMK (Sekolah Menengah Kejuruan) sebanyak 89 sekolah. Pada atribut “Alamat Mahasiswa” yang mendominasi adalah Kabupaten Pemalang sebanyak 181 kecamatan. Dan pada atribut “Program Studi” yang mendominasi adalah S1-Manajemen sebanyak 189 mahasiswa.

Dan jumlah anggota untuk *Cluster 2* sebanyak 186 data. Pada atribut “Jenis Kelamin” yang mendominasi adalah L (Laki-Laki) sebanyak 186 orang. Pada atribut “Asal Sekolah” yang mendominasi adalah SMK (Sekolah Menengah Kejuruan) sebanyak 106 sekolah. Pada atribut “Alamat Mahasiswa” yang mendominasi adalah

Kabupaten Pemalang sebanyak 169 kecamatan. Dan pada atribut “Program Studi” yang mendominasi adalah S1-Manajemen sebanyak 152 mahasiswa.

Sedangkan jumlah anggota untuk *Cluster 3* sebanyak 78 data. Pada atribut “Jenis Kelamin” yang mendominasi adalah P (Perempuan) sebanyak 78 orang. Pada atribut “Asal Sekolah” yang mendominasi adalah SMK (Sekolah Menengah Kejuruan) sebanyak 52 sekolah. Pada atribut “Alamat Mahasiswa” yang mendominasi adalah Kabupaten Pemalang sebanyak 78 kecamatan. Dan pada atribut “Program Studi” yang mendominasi adalah D3-Manajemen Perusahaan sebanyak 40 mahasiswa.

Dari ketiga *cluster* tersebut, diperoleh hasil yang paling mendominasi adalah mahasiswa yang berjenis kelamin perempuan dengan asal Sekolah Menengah Kejuruan dari Kabupaten Pemalang dan memilih program studi S1-Manajemen. Masalah dalam penelitian ini adalah menentukan strategi promosi berdasarkan data Penerimaan Mahasiswa Baru. Disini saya menggunakan atribut daerah asal mahasiswa untuk melakukan pengelompokkan daerah dalam menentukan strategi promosi. Dimana kabupaten yang paling mendominasi adalah Kabupaten Pemalang. Kabupaten tersebut tepat/sama dengan alamat dari Sekolah Tinggi Ilmu Ekonomi Assholeh Pemalang itu sendiri. Hal itu tentunya memudahkan dan memberikan keuntungan bagi *team marketing* karena dapat melakukan promosi secara optimal. Strategi promosi yang dilakukan tentunya akan selaras dengan strategi *promotion mix*. *Team marketing* akan melakukan sosialisasi secara langsung kepada para siswa/siswi untuk mengenalkan Sekolah Tinggi Ilmu Ekonomi Assholeh Pemalang dengan memaparkan visi misi, program studi dan beberapa beasiswa yang diadakan. Membagikan brosur dan menempelkan pamflet sebagai promosi media cetak, melakukan periklanan di sosial media seperti instagram, facebook, twitter dan tentunya website dari Sekolah Tinggi Ilmu Ekonomi Assholeh Pemalang itu sendiri. Selain itu dengan melakukan promosi melalui siaran radio, promosi yang dilakukan dari alumni/alumnus Sekolah Tinggi Ilmu Ekonomi Assholeh Pemalang yang secara tidak langsung akan mempromosikan bahkan mengajak keluarga dan orang-orang terdekat untuk mendaftar di Sekolah Tinggi Ilmu Ekonomi Assholeh Pemalang.

4. KESIMPULAN

Berdasarkan hasil dari penelitian menggunakan algoritma *K-Means Clustering* dari data Penerimaan Mahasiswa Baru pada Tahun Ajaran 2018/2019, 2019/2020, 2020/2021, 2021/2022 di Sekolah Tinggi Ilmu Ekonomi (STIE) Assholeh Pemalang, maka dapat disimpulkan sebagai berikut:

1. Jumlah data yang digunakan 453 *record* dari data Penerimaan Mahasiswa Baru di Sekolah Tinggi Ilmu Ekonomi (STIE) Assholeh Pemalang dan dilakukan uji coba menggunakan *software* WEKA dengan versi 3.8.5 menggunakan algoritma *K-Means Clustering* diperoleh 3 *cluster* yaitu *Cluster 1* berjumlah 189 (42%), *Cluster 2* berjumlah 186 (41%) dan *Cluster 3* berjumlah 78 (17%). Dengan kecepatan komputasi dari perhitungan menggunakan *software* WEKA sebesar 0.01 detik.
2. Dari hasil perhitungan menggunakan persamaan rumus *Euclidean Distance* diperoleh informasi mengenai profil mahasiswa yang mendominasi dalam proses pendaftaran mahasiswa baru tahun ajaran 2018/2019, 2019/2020, 2020/2021 dan 2021/2022 adalah mahasiswa yang berjenis kelamin perempuan dengan asal sekolah yaitu SMK (Sekolah Menengah Kejuruan) yang berasal dari Kabupaten Pemalang dan pada saat mendaftar di Sekolah Tinggi Ilmu Ekonomi Assholeh Pemalang memilih program studi S1-Manajemen.
3. Berdasarkan data hasil keseluruhan proses *clustering* dapat digunakan untuk menentukan strategi promosi yang mampu memudahkan pihak *marketing* Sekolah Tinggi Ilmu Ekonomi (STIE) Assholeh Pemalang supaya proses promosi yang dilakukan lebih efektif dan efisien, sebagai berikut:
 - a. Promosi pada persebaran wilayah di Kabupaten yang mendominasi berdasarkan asal sekolah dari calon mahasiswa baru. Proses ini dilakukan dengan mendatangi secara langsung ke sekolah-sekolah dengan tujuan untuk mengenalkan STIE Assholeh Pemalang dikalangan siswa/siswi terutama kelas XII dengan cara sosialisasi secara langsung dan membagikan brosur serta menempelkan pamflet di papan dinding sekolah sebagai media promosi secara cetak.
 - b. Pihak *marketing* dapat mengirimkan *team* yang memiliki pengetahuan dan potensi lebih dari masing-masing program studi untuk melakukan promosi di kabupaten-kabupaten yang memiliki minat lebih besar pada program studi tersebut dengan cara melakukan penyelarasan menggunakan strategi *promotion mix* yaitu : periklanan, penjualan personal, promosi penjualan, hubungan masyarakat dan pemasaran langsung.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Universitas Stikubank Semarang dan dosen pembimbing atas dukungan dalam menyelesaikan tugas akhir serta terima kasih kepada pihak STIE Assholeh Pemalang yang telah memberikan ijin untuk melakukan penelitian tugas akhir sehingga penulis dapat menyelesaikan tugas akhir dengan baik.

REFERENCES

- [1] Autoridad Nacional del Servicio Civil, "Fungsi Dan Tujuan Pendidikan Indonesia," *Angew. Chemie Int. Ed.* 6(11), 951–952., no. April, pp. 2013–2015, 2021.
- [2] K. Hengki Primayana, "Manajemen Sumber Daya Manusia Dalam Peningkatan Mutu Pendidikan Di Perguruan Tinggi," *J. Penjaminan Mutu*, vol. 1, no. 2, p. 7, 2016, doi: 10.25078/jpm.v1i2.45.
- [3] N. A. Manihuruk, M. Zarlis, E. Irawan, and H. S. Tambunan, "Penerapan Data Mining Dalam Mengelompokkan Calon Penerima Beasiswa Dengan Menggunakan Algoritma K-Means," vol. 4, pp. 29–34, 2020, doi: 10.30865/komik.v4i1.2575.
- [4] F. Nasari and S. Darma, "Seminar Nasional Teknologi Informasi dan Multimedia 2015 PENERAPAN K-MEANS CLUSTERING PADA DATA PENERIMAAN MAHASISWA BARU (STUDI KASUS : UNIVERSITAS POTENSI UTAMA)," pp. 6–8, 2015.
- [5] J. O. Ong, "Implementasi Algoritma K-means clustering untuk menentukan strategi marketing president university," *J. Ilm. Tek. Ind.*, vol. vol.12, no. no. juni, pp. 10–20, 2013.
- [6] T. Arifin, "Penerapan Algoritma K-Means Untuk Pengelompokkan Siswa Lolos SNMPTN Di SMAN 8 Bandung," *J. Responsif*, vol. Vol. 2 No., no. Riset Sains & Informatika, pp. 184–190, 2020.
- [7] S. Bayar, "Implementasi Data Mining Pemilihan Pelanggan Potensial Menggunakan Algoritma K-Means," vol. 1, pp. 72–77, 2018.
- [8] D. Exasanti and A. Jananto, "Analisa Hasil Pengelompokan Wilayah Kejadian Non- Kebakaran Menggunakan Agglomerative Hierarchical Clustering pada Dinas Pemadam Kebakaran Kota Semarang," *J. tekno Kompak*, vol. 15, no. 2, pp. 63–75.
- [9] T. Noviana, J. Jasmir, and Y. Novianto, "Penerapan Data Mining Menentukan Kelompok Prioritas Penerima Bantuan Beras Rastra Dengan Clustering K-Means," *Progr. Stud. Tek. Inform. Stikom Din. Bangsa*, pp. 159–174, 2019.
- [10] A. K. Wardhani, "Implementasi Algoritma K-Means untuk Pengelompokkan Penyakit Pasien pada Puskesmas Kajen Pekalongan," *J. Transform.*, vol. 14, no. 1, pp. 30–37, 2016.
- [11] R. W. Sari, A. Wanto, and A. P. Windarto, "Implementasi Rapidminer Dengan Metode K-Means (Study Kasus: Imunisasi Campak Pada Balita Berdasarkan Provinsi)," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 2, no. 1, pp. 224–230, 2018, doi: 10.30865/komik.v2i1.930.
- [12] E. Sopandi, "STRATEGI PROMOSI PENERIMAAN MAHASISWA BARU UNTUK PTS (Tinjauan atas kegiatan Promosi PTS di Jawa Barat Tahun 2010)," *Manajerial*, vol. 10, no. 1, pp. 87–101, 2011, [Online]. Available: <https://ejournal.upi.edu/index.php/manajerial/article/view/1828>.
- [13] I. Kurniawati, R. E. Indrajit, and M. Fauzi, "Peran Bussines Intelligence Dalam Menentukan Strategi Promosi Penerimaan Mahasiswa Baru," vol. 1, no. 2, pp. 70–79, 2017.
- [14] R. Nainggolan, R. Perangin-Angin, E. Simarmata, and A. F. Tarigan, "Improved the Performance of the K-Means Cluster Using the Sum of Squared Error (SSE) optimized by using the Elbow Method," *J. Phys. Conf. Ser.*, vol. 1361, no. 1, 2019, doi: 10.1088/1742-6596/1361/1/012015.