

# Penentuan Strategi Promosi Universitas Budi Darma Menggunakan Algoritma K-Means Clustering

Natalia Silalahi

Program Studi Teknik Informatika, STMIK Budi Darma, Medan, Indonesia  
Email: natalia.novena.silalahi@gmail.com

**Abstrak**—Universitas Budi Darma merupakan salah satu Perguruan Tinggi Swasta (PTS) di Kota Medan, dimana tidak terlepas dari kegiatan promosi pada setiap tahunnya. Kegiatan promosi yang dilakukan oleh Universitas Budi Darma menggunakan biaya yang cukup besar sehingga promosi yang dilakukan diharapkan juga mendapatkan hasil maksimal. Suksesnya promosi yang dilakukan dapat dilihat dari banyaknya jumlah mahasiswa/I yang mendaftar pada tahun ajaran baru. Promosi yang sukses dapat dilihat dari strategi promosi yang digunakan, penentuan strategi promosi bukanlah pekerjaan yang mudah dikarenakan harus mempertimbangkan beberapa aspek. Untuk membantu menentukan strategi promosi pada Universitas Budi Darma dapat menggunakan data mining, dimana dengan data mining dapat menemukan sebuah pola baru dari kumpulan data yang tersimpan. Salah teknik yang digunakan pada data mining adalah teknik klustering dengan algoritma yang digunakan adalah K-Means. Dengan menggunakan algoritma K-Means hasil terbagi atas 3 kluster dimana pada setiap kluster dapat dilihat hasil untuk penentuan strategi promosi

**Kata Kunci:** Data Mining, Strategi, Promosi, Klustering, K-Means.

**Abstract**—Budi Darma University is one of the Private Universities (PTS) in the city of Medan, which is inseparable from promotional activities every year. Promotional activities carried out by Budi Darma University use a large enough cost so that the promotion carried out is also expected to get maximum results. The success of the promotion can be seen from the large number of students enrolling in the new school year. A successful promotion can be seen from the promotion strategy used, determining the promotion strategy is not an easy job because it must consider several aspects. To help determine the promotion strategy at Budi Darma University can use data mining, where data mining can find a new pattern of stored data sets. One of the techniques used in data mining is clustering technique with the algorithm used is K-Means. By using the K-Means algorithm the results are divided into 3 clusters where in each cluster the results can be seen to determine the promotion strategy

**Keywords:** Data Mining, Strategy, Promotion, Clustering, K-Means.

## 1. PENDAHULUAN

Universitas Budi Darma merupakan salah satu Perguruan Tinggi Swasta (PTS) di Kota Medan. Universitas Budi Darma awal berdiri pada tahun 1996 dimana pada awal berdirinya bernama STMIK Budi Darma dan memiliki 2 (dua) Program Studi yaitu Teknik Informatika (S1) dan Manajemen Informatika (D3). Seiring berjalannya waktu dan sesuai dengan visi juga misi pada STMIK Budi Darma berubah status menjadi Universitas Budi Darma yang dimana sekarang terdapat 6 Program Studi yaitu Teknik Informatika (S1), Teknologi Informasi (S1), Sistem Informasi (S1), Pendidikan Teknologi Informasi (S1), Manajemen Retail (S1) dan Manajemen Informatika (D3) dengan total mahasiswa sebanyak 3200 orang. Universitas Budi Darma tidak terlepas dari kegiatan promosi pada setiap tahunnya, dimana kegiatan promosi dilakukan guna menarik minat bagi mahasiswa/I baru. Didalam kegiatan promosinya Universitas Budi Darma harus bersaing dengan PTS lainnya guna menyaring mahasiswa/I baru untuk mendaftar pada Universitas Budi Darma. Kegiatan promosi yang dilakukan membutuhkan dana yang cukup besar menggunakan media promosi atau bahkan turun langsung ke masyarakat baik didalam kota, luar kota bahkan luar provinsi. Dengan kegiatan promosi yang membutuhkan dana yang cukup besar tersebut diharapkan sesuai dengan hasil yang didapatkan dimana dapat dilihat dari banyaknya mahasiswa/I yang mendaftar pada tahun ajaran baru. Namun didalam menentukan strategi promosi khususnya menentukan lokasi dan tujuan promosi yang tepat bukanlah hal yang mudah, ada banyak factor yang harus diperhatikan sehingga membutuhkan waktu untuk menentukan strategi promosi tersebut. Jumlah mahasiswa/i yang mendaftar berasal dari berbagai daerah dan berasal dari latar belakang lulusan pendidikan yang berbeda baik yang Negeri ataupun Swasta. Hal tersebutlah yang menjadi tolak ukur ataupun pertimbangan untuk menentukan strategi promosi pada asal daerah ataupun sekolah tertentu agar hasil yang didapatkan maksimal bertambah dari pada tahun – tahun sebelumnya.

Untuk menentukan strategi promosi yang digunakan oleh Universitas Budi Darma data mining merupakan sebuah cara yang tepat, dimana data mining mengolah data – data yang telah tersedia sebelumnya khususnya data – data mahasiswa yang ada pada Universitas Budi Darma. Data Mining adalah salah satu cara untuk mencari pengetahuan dan pola dari data dalam jumlah besar. Data mining adalah suatu proses ekstraksi atau penggalian data yang belum diketahui sebelumnya, namun dapat dipahami dan berguna dari database yang besar serta digunakan untuk membuat suatu keputusan bisnis yang sangat penting. Teknik – teknik data mining telah digunakan untuk menemukan pola yang tersembunyi dan memprediksi tren masa depan [1]. Pada data mining terdapat banyak teknik – teknik didalamnya salah satunya adalah Teknik Clustering dan Algoritma K-Means merupakan bagian dari pada Teknik Clustering. Algoritma Clustering K-Means merupakan salah satu algoritma dalam data mining yang berbasis jarak terpendek yang membagi data kedalam beberapa cluster berdasarkan kriteria yang berbasis numerik [2].

Pada penelitian terdahulu yang dilakukan oleh Tria Titiani Chasanah dan Widiyono mengatakan bahwa Clustering K-Means membagi nilai variable yang berbeda – beda. Secara keseluruhan cluster diperoleh berdasarkan dari kota asal dan juga asal jurusan. Penentuan strategi promosi yang dapat ditentukan dengan pendekatan nilai hasil clustering K-Means, pendekatan yang didapat memungkinkan memiliki peluang yang besar untuk meningkatkan jumlah mahasiswa/I baru [3].

Pada penelitian yang dilakukan ini menggunakan Algoritma K-Means diharapkan dapat membantu pihak pengambil keputusan pada Universitas Budi Darma guna menentukan strategi promosi kampus agar hasil yang didapatkan juga maksimal sesuai dengan biaya yang digunakan untuk melakukan kegiatan promosi tersebut.

## 2. METODE PENELITIAN

### 2.1 Data Mining

Data mining merupakan analisis dari peninjauan kumpulan data untuk menemukan hubungan yang tidak diduga dan meringkas data dengan cara berbeda dengan cara yang berbeda dengan sebelumnya, yang dapat dipahami dan bermanfaat bagi pemilik data. Data mining merupakan bidang dari beberapa bidang keilmuan yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistic, database, dan visualisasi untuk penanganan permasalahan pengambilan informasi dari database yang besar. Data Mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terikat dari berbagai database besar. Berdasarkan pengertian data mining yang telah dijelaskan di atas, maka data mining merupakan pengetahuan yang tersembunyi di dalam database yang di proses untuk menemukan pola dan teknik statistik matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi pengetahuan dari database tersebut [4].

### 2.2 Algoritma K-Means

K-means clustering merupakan salah satu metode data clustering non-hirarki yang mengelompokkan data dalam bentuk satu atau lebih cluster/kelompok. Data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu cluster/kelompok dan data yang memiliki karakteristik yang berbeda dikelompokkan dengan cluster/kelompok yang lain sehingga data yang berada dalam satu cluster/kelompok memiliki tingkat variasi yang kecil. Langkah-langkah melakukan clustering dengan metode K-Means adalah sebagai berikut [2]:

- Pilih jumlah cluster k.
- Inisialisasi k pusat cluster ini bisa dilakukan dengan berbagai cara. Namun yang paling sering dilakukan adalah dengan cara random. Pusat-pusat cluster diberiduberi nilai awal dengan angka-angka random.
- Alokasikan semua data/ objek ke cluster terdekat. Kedekatan dua objek ditentukan berdasarkan jarak kedua objek tersebut. Demikian juga kedekatan suatu data ke cluster tertentu ditentukan jarak antara data dengan pusat cluster. Dalam tahap ini perlu dihitung jarak tiap data ke tiap pusat cluster. Jarak paling antara satu data dengan satu cluster tertentu akan menentukan suatu data masuk dalam cluster mana. Untuk menghitung jarak semua data ke setiap titik pusat cluster dapat menggunakan teori jarak Euclidean yang dirumuskan sebagai berikut:

$$D(i,j) = \sqrt{(X_{1i} - Y_{1j})^2 + (X_{2i} - Y_{2j})^2 + \dots + (X_{ki} - Y_{kj})^2}$$

Dimana

D (i,j) = Jarak data ke i ke pusat cluster j

X<sub>ki</sub> = Data ke i pada atribut data ke k

X<sub>kj</sub> = Titik pusat ke j pada atribut ke k

- Hitung kembali pusat cluster dengan keanggotaan cluster yang sekarang. Pusat cluster adalah rata – rata dari semua data/ objek dalam cluster tertentu. Jika dikehendaki bisa juga menggunakan median dari cluster tersebut. Jadi rata – rata (mean) bukan satu – satunya ukuran yang bisa di pakai.
- Tugaskan lagi setiap objek memakai ckuster yang bary. Jika pusat cluster tidak berubah lagi maka proses clustering selesai. Atau, kembali ke langkah nomor 3 sampai pusat cluster tidak berubah lagi.

### 2.3 Promosi

Promosi berasal dari kata bahasa Inggris promote yang berarti “meningkatkan” atau “mengembangkan”. Pengertian tersebut jika digunakan dalam bidang penjualan berarti alat untuk meningkatkan omzet penjualan. Pengertian promosi dapat dipandang berbeda dalam hal produsen dan konsumen [5].

## 3. HASIL DAN PEMBAHASAN

### 3.1 Analisa Masalah

Universitas Budi Darma merupakan salah satu Perguruan Tinggi Swasta (PTS) dikota Medan yang berkembang saat ini dengan memiliki jumlah mahasiswa sebanyak 3200 orang, dimana terdapat 6 program studi pada Universitas Budi Darma. Kegiatan promosi merupakan kegiatan rutin yang dilakukan setiap tahun oleh PTS dikota medan, tidak terkecuali Universitas Budi Darma juga melakukan kegiatan promosi tersebut. Kegiatan promosi dilakukan guna menarik minat bagi mahasiswa/I baru baik dikota Medan, Luar Kota bahkan Luar Provinsi untuk masuk dan mendaftar di Universitas Budi Darma. Kegiatan promosi itu sendiri menggunakan biaya yang cukup besar setiap tahunnya dengan menggunakan berbagai media dan bahkan turun langsung ke daerah ataupun sekolah – sekolah. Dikarenakan biaya pada kegiatan promosi cukup besar sehingga diharapkan hasil yang maksimal dengan banyaknya mahasiswa/I baru mendaftar pada

tahun ajaran baru. Untuk mendapatkan hasil promosi yang maksimal maka diperlukannya strategi promosi yang tepat pula, tetapi penentuan strategi promosi bukanlah hal yang mudah udah didapatkan. Penentuan strategi promosi dapat dilihat dari jumlah mahasiswa/I dari asal daerah ataupun dari latar belakang pendidikan ketika SMA, hal tersebutlah yang menjadi tolak ukur bagi pengambil keputusan untuk menentukan strategi promosi di Universitas Budi Darma. Dimana untuk menyelesaikan permasalahan penentuan strategi promosi dapat menerapkan teknik data mining. Data mining itu sendiri merupakan sebuah proses penemuan pola ataupun informasi baru dari kumpulan – kumpulan data dari sebuah organisasi juga perusahaan, dimana pola ataupun informasi baru tersebut digunakan sebagai acuan atau dasar didalam pengambilan keputusan. Salah satu teknik pada data mining yang sering digunakan adalah teknik clustering dimana salah satau algoritmanya adalah K-Means. Algoritma K-Means berbasis jarak terpendek untuk menghasilkan informasi ataupun pola baru, pada K-Means hasil terbagi atas beberapa cluster yang berbeda. Dengan dilakukannya penelitian ini diharapkan menemukan pola strategi promosi yang tepat bagi Universitas Budi Darma.

**3.2 Penerapan Algoritma K-Means.**

Untuk dapat menentukan strategi promosi yang tepat bagi Universitas Budi Darma berdasarkan jumlah mahasiswa/I dari setiap daerah dan juga latar belakang pendidikan ketika SMA. Pada penelitian ini menggunakan data sampling seperti terlihat pada table – table berikut.

**Tabel 1.** Asal Daerah Mahasiswa

| No | Asal Daerah          | Jumlah Mahasiswa | Inisialisasi |
|----|----------------------|------------------|--------------|
| 1  | Medan                | 574              | 1            |
| 2  | Deli Serdang         | 403              | 2            |
| 3  | Serdang Berdagai     | 286              | 3            |
| 4  | Nias                 | 275              | 4            |
| 5  | Aceh                 | 188              | 5            |
| 6  | Tapanuli Tengah      | 143              | 6            |
| 7  | Labuhan Batu Utara   | 123              | 7            |
| 8  | Kisaran              | 122              | 8            |
| 9  | Simalungun           | 120              | 9            |
| 10 | Kota Lainnya         | 119              | 10           |
| 11 | Tapanuli Utara       | 116              | 11           |
| 12 | Tebing Tinggi        | 113              | 12           |
| 13 | Labuhan Batu Selatan | 113              | 13           |
| 14 | Rantau Prapat        | 112              | 14           |
| 15 | Padangsidempuan      | 111              | 15           |
| 16 | Mandailing Natal     | 99               | 16           |
| 17 | Tapanuli Selatan     | 98               | 17           |
| 18 | Riau                 | 85               | 18           |

**Tabel 2.** Asal Sekolah Mahasiswa

| No | Asal Sekolah    | Jumlah Mahasiswa | Inisialisasi |
|----|-----------------|------------------|--------------|
| 1  | SMA Swasta IPA  | 555              | 1            |
| 2  | SMA Swasta IPS  | 440              | 2            |
| 3  | SMA Negeri IPA  | 380              | 3            |
| 4  | SMA Negeri IPS  | 340              | 4            |
| 5  | SMK Swasta RPL  | 335              | 5            |
| 6  | SMK Swasta TKJ  | 310              | 6            |
| 7  | Jurusan Lainnya | 275              | 7            |
| 8  | SMK Negeri RPL  | 180              | 8            |
| 9  | SMK Negeri TKJ  | 140              | 9            |
| 10 | MA Negeri       | 135              | 10           |
| 11 | MA Swasta       | 110              | 11           |

**Tabel 3.** Pekerjaan Orang Tua Mahasiswa

| No | Pekerjaan Orang Tua | Jumlah Mahasiswa | Inisialisasi |
|----|---------------------|------------------|--------------|
| 1  | Wiraswasta          | 695              | 1            |
| 2  | Nelayan             | 660              | 2            |
| 3  | Petani              | 650              | 3            |
| 4  | Buruh               | 535              | 4            |
| 5  | Karyawan Swasta     | 485              | 5            |
| 6  | Pegawai Negeri      | 120              | 6            |
| 7  | Pekerjaan Lainnya   | 55               | 7            |

Tabel 4. Data Sampling

| No | Sampling | Daerah | Sekolah | Pekerjaan |
|----|----------|--------|---------|-----------|
| 1  | S1       | 1      | 1       | 1         |
| 2  | S2       | 2      | 1       | 3         |
| 3  | S3       | 4      | 2       | 2         |
| 4  | S4       | 1      | 2       | 5         |
| 5  | S5       | 6      | 6       | 3         |
| 6  | S6       | 4      | 2       | 3         |
| 7  | S7       | 3      | 1       | 1         |
| 8  | S8       | 5      | 10      | 1         |
| 9  | S9       | 10     | 6       | 2         |
| 10 | S10      | 8      | 6       | 5         |
| 11 | S11      | 2      | 5       | 5         |
| 12 | S12      | 2      | 4       | 7         |
| 13 | S13      | 1      | 2       | 4         |
| 14 | S14      | 1      | 3       | 6         |
| 15 | S15      | 1      | 2       | 5         |

Iterasi – 1

Penentuan nilai *Centeroid* / titik pusat awal yang digunakan pada penelitian diambil secara random dari data sampling.

Tabel 5. *Centeroid* / Titik Pusat Awal

| No | Sampling | Daerah | Sekolah | Pekerjaan | Cluster |
|----|----------|--------|---------|-----------|---------|
| 1  | 4        | 1      | 2       | 5         | C1      |
| 2  | 8        | 5      | 10      | 1         | C2      |
| 3  | 12       | 2      | 4       | 7         | C3      |

Untuk mengetahui *cluster* mana yang paling dekat dengan data, maka perlu dihitung jarak setiap data dengan titik pusat setiap *cluster*. akan dihitung jarak dari data santri pertama ke pusat *cluster* pertama, menggunakan teori jarak *Euclidean*.

$$D(i,j) = \sqrt{(X_{1i} - Y_{1j})^2 + (X_{2i} - Y_{2j})^2 + \dots + (X_{ki} - Y_{kj})^2}$$

a. Menghitung Jarak dari data ke – 1 sampai data ke – 15 dengan *centeroid* pertama (C1)

$$D(1,1) = \sqrt{(1 - 1)^2 + (1 - 2)^2 + (1 - 5)^2} = 4,123106$$

$$D(2,1) = \sqrt{(2 - 1)^2 + (1 - 2)^2 + (3 - 5)^2} = 2,44949$$

$$D(3,1) = \sqrt{(4 - 1)^2 + (2 - 2)^2 + (2 - 5)^2} = 4,242641$$

$$D(4,1) = \sqrt{(1 - 1)^2 + (2 - 2)^2 + (5 - 5)^2} = 0$$

$$D(5,1) = \sqrt{(6 - 1)^2 + (6 - 2)^2 + (3 - 5)^2} = 6,708204$$

b. Menghitung Jarak dari data ke – 1 sampai data ke – 15 dengan *centeroid* kedua (C2)

$$D(1,1) = \sqrt{(1 - 1)^2 + (1 - 2)^2 + (1 - 5)^2} = 9,8488578$$

$$D(2,1) = \sqrt{(2 - 1)^2 + (1 - 2)^2 + (3 - 5)^2} = 9,6953597$$

$$D(3,1) = \sqrt{(4 - 1)^2 + (2 - 2)^2 + (2 - 5)^2} = 8,1240384$$

$$D(4,1) = \sqrt{(1 - 1)^2 + (2 - 2)^2 + (5 - 5)^2} = 9,797959$$

$$D(5,1) = \sqrt{(6 - 1)^2 + (6 - 2)^2 + (3 - 5)^2} = 4,5825757$$

c. Menghitung Jarak dari data ke – 1 sampai data ke – 15 dengan *centeroid* ketiga (C3)

$$D(1,1) = \sqrt{(1 - 1)^2 + (1 - 2)^2 + (1 - 5)^2} = 6,78233$$

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

$$D(3,1) = \sqrt{(4 - 1)^2 + (2 - 2)^2 + (2 - 5)^2} = 5,7445626$$

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

$$D(5,1) = \sqrt{(6 - 1)^2 + (6 - 2)^2 + (3 - 5)^2} = 6$$

Adapun hasil perhitungan jarak data ke – 1 sampai data ke – 15 terhadap *centeroid* / titik pusat awal C1, C2 dan C3 dapat dilihat pada tabel 6.

Tabel 6. Hasil Perhitungan Jarak

| No | Sampling | C1        | C2        | C3      | Hasil |
|----|----------|-----------|-----------|---------|-------|
| 1  | S1       | 4,1231056 | 9,8488578 | 6,78233 | 1     |
| 2  | S2       | 2,4494897 | 9,6953597 | 5       | 1     |

|    |     |           |           |           |   |
|----|-----|-----------|-----------|-----------|---|
| 3  | S3  | 4,2426407 | 8,1240384 | 5,7445626 | 1 |
| 4  | S4  | 0         | 9,797959  | 3         | 1 |
| 5  | S5  | 6,7082039 | 4,5825757 | 6         | 2 |
| 6  | S6  | 3,6055513 | 8,3066239 | 4,8989795 | 1 |
| 7  | S7  | 4,5825757 | 9,2195445 | 6,78233   | 1 |
| 8  | S8  | 9,797959  | 0         | 9         | 2 |
| 9  | S9  | 10,29563  | 6,4807407 | 9,6436508 | 2 |
| 10 | S10 | 8,0622577 | 6,4031242 | 6,6332496 | 2 |
| 11 | S11 | 3,1622777 | 7,0710678 | 2,236068  | 3 |
| 12 | S12 | 3         | 9         | 0         | 3 |
| 13 | S13 | 1         | 9,4339811 | 3,7416574 | 1 |
| 14 | S14 | 1,4142136 | 9,486833  | 1,7320508 | 1 |
| 15 | S15 | 0         | 9,797959  | 3         | 1 |

**Iterasi - 2**

Setelah semua data ditempatkan ke dalam cluster yang terdekat. Hitung kembali pusat cluster dengan keanggotaan cluster yang sekarang. Pusat cluster adalah rata-rata dari semua data/obyek dalam cluster tertentu. bisa juga memakai median dari cluster tersebut. Jadi rata-rata (mean) bukan satu-satunya ukuran yang dapat digunakan. Hitung pusat cluster baru.

a. Untuk *cluster* 1 terdapat 9 data.

$$K_{11} = \frac{4,1231056+2,4494897+4,2426407+0+3,6055513+4,5825757+1+1,4142136+0}{9} = \frac{21,41758}{9} = 2,379731$$

$$K_{12} = \frac{9,8488578+9,6953597+8,1240384+9,797959+8,3066239+9,2195445+9,4339811+9,486833+9,797959}{9} = \frac{83,7116}{9} = 9,30124$$

$$K_{13} = \frac{6,78233+5+5,7445626+3+4,8989795+6,78233+3,7416574+1,7320508+3}{9} = \frac{40,68191}{9} = 4,520212$$

b. Untuk *cluster* 2 terdapat 4 data.

$$K_{21} = \frac{6,7082039+9,797959+10,29563+8,0622577}{4} = \frac{34,86405}{4} = 8,716013$$

$$K_{22} = \frac{4,5825757+0+6,4807407+6,4031242}{4} = \frac{17,46644}{4} = 4,36661$$

$$K_{23} = \frac{6+9+9,6436508+6,6332496}{4} = \frac{31,2769}{4} = 7,819225$$

c. Untuk *cluster* 3 terdapat 3 data.

$$K_{31} = \frac{3,1622777+3}{2} = \frac{6,162278}{2} = 3,081139$$

$$K_{32} = \frac{7,0710678+9}{2} = \frac{16,07107}{2} = 8,035534$$

$$K_{33} = \frac{2,236068+0}{2} = \frac{2,236068}{2} = 1,118034$$

**Tabel 7.** *Centeroid* / Titik Pusat Awal Iterasi – 2

| No | Cluster | K1       | K2       | K3       |
|----|---------|----------|----------|----------|
| 1  | C1      | 2,379731 | 9,30124  | 4,520212 |
| 2  | C2      | 8,716013 | 4,36661  | 7,819225 |
| 3  | C3      | 3,081139 | 8,035534 | 1,118034 |

Setelah didapatkan *centeroid* / titik pusat awal untuk C1, C2 dan C3 pada iterasi kedua, selanjutnya menghitung jarak kembali dari data ke – 1 sampai data ke – 15 dengan jarak *centeroid* pada iterasi – 2. Adapun hasil perhitungan jarak pada iterasi – 2 dengan model *Euclidean* adalah sebagai berikut :

**Tabel 8.** Hasil Perhitungan Jarak Iterasi – 2

| No | Samplng | C1        | C2        | C3        | Hasil |
|----|---------|-----------|-----------|-----------|-------|
| 1  | S1      | 9,1217397 | 10,83387  | 7,3378342 | 3     |
| 2  | S2      | 8,4478296 | 8,9254593 | 7,3627031 | 3     |
| 3  | S3      | 7,8920749 | 7,8552531 | 6,1684553 | 3     |
| 4  | S4      | 7,4459358 | 8,5490192 | 7,4718452 | 1     |
| 5  | S5      | 5,1298711 | 5,7679822 | 4,0255365 | 3     |
| 6  | S6      | 7,6318033 | 7,1460864 | 6,3885657 | 3     |
| 7  | S7      | 9,0380978 | 9,5136059 | 7,0369918 | 3     |
| 8  | S8      | 4,4436435 | 9,5940432 | 2,7486517 | 3     |
| 9  | S9      | 8,6786032 | 6,1789939 | 7,2658035 | 2     |
| 10 | S10     | 6,535733  | 3,3359657 | 6,5884939 | 2     |
| 11 | S11     | 4,3445434 | 7,3112272 | 5,0450955 | 1     |
| 12 | S12     | 5,8648691 | 6,7757186 | 7,2147017 | 1     |
| 13 | S13     | 7,4486498 | 8,9288397 | 7,0046084 | 3     |

| No | Sampling | C1        | C2        | C3        | Hasil |
|----|----------|-----------|-----------|-----------|-------|
| 14 | S14      | 6,6180855 | 8,0445049 | 7,3158277 | 1     |
| 15 | S15      | 7,4459358 | 8,5490192 | 7,4718452 | 1     |

**Iterasi – 3**

Dikarenakan terdapat perbedaan hasil clustering padang Iterasi – 1 dan Iterasi – 2 maka proses dilanjutkan pada iterasi 3 dengan cara yang sama. Untuk *Centeroid* / titik pusat iterasi – 3 dapat dilihat pada tabel berikut.

**Tabel 9.** *Centeroid* / Titik Pusat Iterasi – 3

| No | Cluster | K1       | K2       | K3       |
|----|---------|----------|----------|----------|
| 1  | C1      | 6,343874 | 7,845898 | 6,903863 |
| 2  | C2      | 7,607168 | 4,75748  | 6,927149 |
| 3  | C3      | 3,697107 | 4,285321 | 3,004584 |

**Tabel 10.** Hasil Perhitungan Jarak Iterasi – 3

| No | Sampling | C1        | C2        | C3        | Hasil |
|----|----------|-----------|-----------|-----------|-------|
| 1  | S1       | 10,501376 | 9,6386939 | 4,6995827 | 3     |
| 2  | S2       | 8,9986503 | 7,8090645 | 3,6977733 | 3     |
| 3  | S3       | 7,9822391 | 6,7001608 | 2,5146819 | 3     |
| 4  | S4       | 3,987765  | 7,4143286 | 4,0594043 | 1     |
| 5  | S5       | 4,3319436 | 4,421464  | 2,8711428 | 3     |
| 6  | S6       | 7,4100213 | 6,0031539 | 2,3053107 | 3     |
| 7  | S7       | 9,6386416 | 8,3949835 | 3,9112209 | 3     |
| 8  | S8       | 6,4266438 | 8,3314126 | 6,1946302 | 3     |
| 9  | S9       | 6,3892464 | 5,6166091 | 6,6087649 | 2     |
| 10 | S10      | 3,1264656 | 2,3263869 | 5,0434807 | 2     |
| 11 | S11      | 5,8311003 | 5,9340587 | 5,7152575 | 1     |
| 12 | S12      | 5,3025352 | 5,6585702 | 5,8502792 | 1     |
| 13 | S13      | 8,4358718 | 7,7347635 | 3,6725919 | 3     |
| 14 | S14      | 5,2702603 | 6,8994935 | 6,2307155 | 1     |
| 15 | S15      | 6,1459319 | 7,4143286 | 7,0594043 | 1     |

Pada iterasi – 3 tidak ada perbedaan hasil dengan iterasi – 2. Oleh sebab itu, perhitungan jarak terdekat berhenti. Maka dapat dikelompokkan sebagai berikut:

**Tabel 11.** Cluster 1

| No | Sampling | C1        | C2        | C3        | Hasil |
|----|----------|-----------|-----------|-----------|-------|
| 4  | S4       | 3,987765  | 7,4143286 | 4,0594043 | 1     |
| 11 | S11      | 5,8311003 | 5,9340587 | 5,7152575 | 1     |
| 12 | S12      | 5,3025352 | 5,6585702 | 5,8502792 | 1     |
| 14 | S14      | 5,2702603 | 6,8994935 | 6,2307155 | 1     |
| 15 | S15      | 6,1459319 | 7,4143286 | 7,0594043 | 1     |

**Tabel 12.** Cluster 2

| No | Sampling | C1        | C2        | C3        | Hasil |
|----|----------|-----------|-----------|-----------|-------|
| 9  | S9       | 6,3892464 | 5,6166091 | 6,6087649 | 2     |
| 10 | S10      | 3,1264656 | 2,3263869 | 5,0434807 | 2     |

**Tabel 13.** Cluster 3

| No | Sampling | C1        | C2        | C3        | Hasil |
|----|----------|-----------|-----------|-----------|-------|
| 1  | S1       | 10,501376 | 9,6386939 | 4,6995827 | 3     |
| 2  | S2       | 8,9986503 | 7,8090645 | 3,6977733 | 3     |
| 3  | S3       | 7,9822391 | 6,7001608 | 2,5146819 | 3     |
| 5  | S5       | 4,3319436 | 4,421464  | 2,8711428 | 3     |
| 6  | S6       | 7,4100213 | 6,0031539 | 2,3053107 | 3     |
| 7  | S7       | 9,6386416 | 8,3949835 | 3,9112209 | 3     |
| 8  | S8       | 6,4266438 | 8,3314126 | 6,1946302 | 3     |
| 13 | S13      | 8,4358718 | 7,7347635 | 3,6725919 | 3     |

Berdasarkan hasil proses clustering menggunakan K-Means maka didapatkan sebuah informasi untuk penentuan promosi sebagai berikut :

1. *Cluster 1 (C1)*

Kluster 1 didominasi asal daerah dari Kota Medan, Untuk latar belakang pendidikan sekolah berasal SMA Swasta dengan penjurusan IPS dan pekerjaan orang tua adalah karyawan swasta

2. *Cluster 2 (C2)*

Kluster 2 asal daerah dari Kisaran, latar belakang pendidikan sekolah berasal dari SMK Swasta dengan penjurusan TKJ dan pekerjaan orang tua adalah karyawan swasta.

3. *Cluster 3 (C3)*

Kluster 3 didominasi asal daerah dari Kota Medan juga Nias, latar belakang pendidikan sekolah berasal dari SMA Swasta dengan penjurusan IPA juga IPS dan pekerjaan orang tua adalah wiraswasta dan petani.

## 4. KESIMPULAN

Berdasarkan uraian penelitian diatas, maka dapat ditarik sebuah kesimpulan sebagai berikut :

1. Algoritma K-Means dapat dipergunakan untuk melakukan clustering guna menentukan strategi promosi pada Universitas Budi Darma.
2. Dengan menggunakan Algoritma K-Means clustering terbagi atas 3 kelompok.
3. Dari hasil clustering didapatkan bahwa untuk daerah asal yang menjadi target promosi adalah Kota Medan dan latar belakang pendidikan sekolah SMA Swasta dengan penjurusan IPA dan IPS.

## DAFTAR PUSTAKA

- [1] E. H. Pujiarini, "ANALISIS ASOSIASI UNTUK MENENTUKAN STRATEGI PROMOSI PERGURUAN TINGGI DENGAN ALGORITMA APRIORI," Jurnal Informatika dan Komputer (JIKO), vol. 4, no. 1, pp. 45-51, 2019.
- [2] E. Buulolo dan R. Syahputra, "Implementasi Algoritma Clustering K-Means Untuk Mengelompokkan Mahasiswa Baru Yang Berpotensi (Studi Kasus: Stmik Budi Darma)," dalam Seminar Nasional Riset Information Science (SENARIS), Siantar, 2019.
- [3] T. T. Chasanah dan Widiyono, "PENENTUAN STRATEGI PROMOSI PENERIMAAN MAHASISWA BARU DENGAN ALGORITMA CLUSTERING K-MEANS," IC-Tech, vol. XII, no. 1, pp. 39-44, 2017.
- [4] D. P. Utomo dan Mesran, "Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung," JURNAL MEDIA INFORMATIKA BUDIDARMA, vol. 4, no. 2, pp. 437-444, 2020.
- [5] S. Setiawan, "GURUPENDIDIKAN.COM," 28 Januari 2020. [Online]. Available: <https://www.gurupendidikan.co.id/pengertian-promosi/>. [Diakses 17 Juni 2020].