

Penerapan Algoritme *K-Means Cluster* dan Metode TOPSIS pada Pemilihan Mahasiswa Kunjungan Industri

Stendy Budi Hartono Sakur^{1*}, Miske Silangen², Desmin Tuwohingide³

Program Studi Sistem Informasi, Politeknik Negeri Nusa Utara Tahuna

Jl. Kesehatan No.1 Sawang Bendar, Tahuna, Indonesia

*e-mail *Corresponding Author*: sakur.stendy@gmail.com

Abstract

Industrial visits are an applicable form of learning where educators and students can see directly the system or work pattern of an industry or software company. The importance of this activity, so it is necessary to select students who take part in it accurately and objectively. This study aims to make calculation standards to help facilitate study programs in selecting students. Using the K-Means Cluster Algorithm is used to streamline data by eliminating alternatives that do not meet the requirements and creating two clusters whose centroid initially uses the maximum and minimum values of the criteria. Then do the ranking process with the TOPSIS method. The results showed that calculating the distance using the Manhattan Distance has the closest coefficient value higher than the Euclidean distance, which is about 60% of all data. The first cluster consists of 25 people who meet the requirements and are ranked by the TOPSIS Method so that only 20 people are left to take part in the activity. By using the Euclidean distance, there are 70% of the 20 people selected, while the Manhattan distance is 75%. 30% and 25% are taken from the first and second clusters. The K-Means Algorithm can correctly group members according to the required characteristics so that it can streamline the initial data, then the MCDM method can speed up the calculation process accurately and objectively.

Keywords: *K-Means; Clustering; Metode TOPSIS; Manhattan distance; Eucliden distance*

Abstrak

Kunjungan industri merupakan salah satu bentuk pembelajaran yang aplikatif dimana pendidik dan peserta didik dapat melihat secara langsung sistem atau pola kerja suatu industri atau perusahaan perangkat lunak. Pentingnya kegiatan ini, sehingga perlu untuk memilih mahasiswa yang mengambil bagian di dalamnya secara akurat dan obyektif. Penelitian ini bertujuan membuat standar perhitungan untuk membantu mempermudah program studi dalam memilih mahasiswa. Menggunakan Algoritme *K-Means Cluster* yang digunakan untuk merampingkan data dengan Cara mengeliminasi alternatif yang tidak memenuhi syarat serta membuat dua kluster yang centroid awalnya menggunakan nilai maksimum dan minimum dari kriteria. Kemudian lakukan proses perankingan dengan metode TOPSIS. Hasil penelitian menunjukkan, perhitungan jarak dengan *Manhattan Distance* memiliki nilai koefisien terdekat lebih tinggi dari *Euclidean distance* sekitar 60% dari seluruh data. *Cluster* pertama terdiri dari 25 orang yang memenuhi persyaratan dan diranking dengan Metode TOPSIS, sehingga tersisa 20 orang untuk mengikuti kegiatan tersebut. Dengan menggunakan *Eucliden distance*, terdapat 70% dari 20 orang yang dipilih, sedangkan *Manhattan distance* adalah 75%. 30% dan 25% diambil dari *cluster* pertama dan kedua. Algoritme *K-Means* dapat dengan tepat mengelompokkan anggota sesuai karakteristik yang diperlukan sehingga dapat merampingkan data awal, kemudian metode MCDM dapat mempercepat proses perhitungan dengan akurat serta objektif.

Kata kunci: *K-Means; Clustering; Metode TOPSIS; Jarak Manhattan; Jarak Eucliden*

1. Pendahuluan

Kemampuan dalam pengembangan diri bagi mahasiswa dapat diwujudkan melalui berbagai macam kegiatan, salah satu diantaranya adalah melakukan pembelajaran secara langsung dengan melihat bagaimana system dan pola kerja dari sebuah perusahaan yang erat hubungannya dengan program studi yang ditempuh oleh mahasiswa atau dikenal dengan kegiatan kunjungan industri. Tujuan dari kegiatan ini adalah memberikan wawasan kepada

mahasiswa tentang proses kerja yang ada disektor industri khususnya bidang perangkat lunak yang dapat memberikan perubahan terhadap mutu kurikulum, perubahan system serta kebijakan lainnya dalam proses peningkatan kualitas lulusan dari pendidikan vokasi. Kegiatan ini dapat membantu pendidik serta peserta didik tentang gambaran yang lebih luas dari kompetensi yang dibutuhkan oleh dunia industri sehingga proses pembekalan telah dilakukan selama dalam perkuliahan untuk mempersiapkan mahasiswa memasuki dunia kerja sehingga penyerapan tenaga kerja pendidikan vokasi lebih cepat dan tepat sasaran. Kolaborasi antara dunia industri dan pendidikan merupakan kunci keberhasilan dari pendidikan vokasi karena dapat memberikan perubahan yang signifikan terhadap kompetensi dan standar kemampuan dari peserta didik [1]. Selain manfaat kegiatan ini terhadap perubahan dari kualitas pendidik dan peserta didik, juga memberikan perubahan terhadap pihak perusahaan manufaktur dan industry perangkat lunak karena pihak perusahaan dapat memantua sejauh mana kesiapan dari perusahaan dalam memberikan pelayanan yang terbaik selama kegiatan tersebut berlangsung melalui feedback dari pengunjung [2]. Dengan adanya kegiatan ini, maka setiap perusahaan manufaktur ataupun industry perangkat lunak telah membuat dan menerapkan prosedur ataupun tatacara pengajuan kegiatan kunjungan industri.

Politeknik Negeri Nusa Utara merupakan salah satu perguruan tinggi vokasi yang ada diujung utara Indonesia dan merupakan kebanggaan dari daerah Kabupaten kepulauan sangihe. Salah satu program studi yang banyak diminati adalah Sistem informasi yang setiap tahunnya mahasiswa diberikan kesempatan untuk dapat secara nyata melihat sistem kerja pada industri perangkat lunak. Didaerah Kabupaten Kepulauan Sangihe itu sendiri belum terdapat industri perangkat lunak yang berkembang pesat sehingga harus dilakukan diluar daerah dengan peserta yang dibatasi. Dengan demikian perlu dilakukan pemilihan mahasiswa yang memenuhi persyaratan berdasarkan kriteria tertentu melalui suatu mekanisme pengelompokan ataupun cluster yang anggotanya dirangking dengan mekanisme pendukung keputusan sehingga dapat membantu pimpinan dalam pengambilan keputusan.

Penelitian terhadulu tentang pola pemikiran ini diantaranya Hastuti [3] yang menggunakan *Fuzzy C-Means* dalam proses cluster dan menggunakan TOPSIS untuk menentukan jurusan di SMA, namun tidak mempertimbangkan pengaruh pengukuran jarak dari metode TOPSIS, Budiana [4] menggunakan K-Means dan Topsis untuk perangkingan dengan eucliden distance pada penetapan infrastruktur diklat pada PT.PLN (persero), Mirfan [5] dengan aplikasi berbasis web membuat solusi pada pemilihan daerah wisata menggunakan algoritma K-Means dan perangkingan menggunakan Topsis, Dainati [6] menggunakan K-Means untuk meningkatkan kemampuan dalam pengambilan keputusan menggunakan metode SAW dan TOPSIS, Khomsatun [7] dalam penelitiannya menggabungkan K-Means Clustering dengan TOPSIS untuk menentukan pemilihan lahan tanam, Masruro [8] melakukan penelitian sama halnya dengan Mirfan untuk menentukan lokasi wisata dengan proses clustering dengan K-Means dan perangkingan menggunakan TOPSIS, Raharja [9] meneliti tentang strategi pemasaran menggunakan algoritma K-Means dan metode TOPSIS, Sun [10] mengevaluasi dan merangking kinerja perkantoran menggunakan *Entropy-based TOPSIS* dan *K-Means*, Trstenjak [11] menggunakan *K-Means* dan *Fuzzy TOPSIS* untuk mengevaluasi strategi pengembangan Kroasia dengan Analisis SWOT, kemudian Wiyantoro melakukan analisis penyerang terbaik Liga Inggris dengan menggunakan rasio *Fisher's Discriminant Ratio* beserta *K-Means* dan TOPSIS, dalam kasus ini *Fisher's Discriminant* digunakan untuk menganalisa variabel yang kontinu dan binary selanjutnya akan di *cluster* oleh K-Means dan dirangking oleh TOPSIS.

Dalam kaitannya dengan pengukuran jarak baik dalam *K-Means* maupun TOPSIS terdapat penelitian terdahulu semisal, Haviluddin [12] melakukan perbandingan jarak *Eucliden*, *Manhattan* dan *Minkowski distance* pada *K-Means* untuk mengcluster tingkat kejahatan yang menunjukkan ketiga jarak tersebut memberikan hasil yang ideal, demikian pula Nishom [13] yang membandingkan *Eucliden*, *Minkowski*, dan *Manhattan distance* pada algoritme *K-Means* dengan memperhitungkan nilai *Chi-Square* yang hasilnya menunjukkan ketiganya memiliki akurasi yang baik, Setiawan [14] melakukan perbandingan jarak *Manhattan*, *Euclidean* dan *Minkowski distance* dalam proses klasifikasi data Iris pada Metode KNN dimana perubahan yang cukup besar kepada *Minkowski distance* jika nilai p mengalami perubahan, selain itu penelitian dari Chiu [15] menganalisa minimum Manhattan distance dalam MCDM untuk *Multiobjective Optimization Problems* (MOPs).

Dari penelitian terhadulu terlihat bahwa penelitian yang menggabungkan Algoritme *K-Means* dan TOPSIS tidak mempertimbangkan masalah perubahan yang terjadi pada penentuan

nilai ideal positif dan negatif. Oleh karena itu penelitian ini bertujuan untuk melihat pengaruh dari pengukuran jarak *Euclidean Distance* dan *Manhattan Distance* pada Metode TOPSIS untuk penentuan nilai solusi ideal positif dan negatif yang datanya telah dirampingkan atau dikelompokkan menggunakan Algoritme *K-Means*. Sebagai pembandingan digunakan data mahasiswa yang telah selesai melaksanakan kegiatan tersebut untuk melihat ketepatan atau akurasi dari perhitungan TOPSIS.

2. Metodologi

Skenario yang digunakan pada penelitian ini adalah melakukan perampingan atau pengelompokan data awal dengan Algoritme *K-Means* sehingga hanya data yang memiliki karakteristik sesuai kriteria yang dibutuhkan untuk diproses. Kluster dibagi menjadi 2 bagian dalam hal ini nilai k, dimana kluster pertama untuk anggota dengan karakteristik yang memenuhi syarat sedangkan kluster lainnya yang tidak memenuhi. Nilai *centroid* awal menggunakan nilai Maksimum dan Minimum untuk karakteristik yang memenuhi syarat dan tidak. Setelah proses clustering dilakukan maka anggota pada cluster pertama akan diranking menggunakan metode MCDM yaitu TOPSIS. Dalam proses perankingan, metode TOPSIS akan menghitung nilai Ideal positif dan negatif menggunakan *Euclidean distance (default dari Topsis)* dan *Manhattan distance*. Selanjutnya hasil perankingan dibandingkan dengan data mahasiswa yang telah melaksanakan kegiatan tersebut.

2.1 Sumber Data

Data yang digunakan pada penelitian ini meminjam data dari penelitian [16] yang dapat didownload di repositori [17]. Data tersebut diambil menggunakan kuisioner yang dilakukan pada tahun 2019 dari mahasiswa angkatan 2017 dengan jumlah responden sebanyak 55 orang.

2.2. Pre-Processing

Data awal yang diperoleh akan dimodifikasi dengan mengisi nilai tertentu pada data yang kosong atau nol agar proses dapat dilakukan dengan benar. Kemudian, data awal tersebut harus dikonversikan kedalam nilai numerik untuk beberapa kriteria yang menggunakan nilai linguistik. Tabel 1, menunjukkan batasan konversi ke nilai numerik.

Tabel 1. Nilai Konversi Numerik

Kriteria	Linguistik	Numerik
Tes Wawancara	Memuaskan	1
	Cukup	0.8
	Kurang	0.6
	Tidak ada	0.4
Keaktifan di kelas	Aktif	1
	Sedikit Aktif	0.6
	Kurang Aktif	0.2
Kepemimpinan	Ketua	1
	Sekretaris	0.8
	Bendahara	0.6
	Seksi	0.4
	Tidak ada	0.1
Unit Kegiatan Mahasiswa (UKM)	Aktif	1
	Kurang Aktif	0.6
	Tidak Aktif	0.2
Sertifikat Keahlian	Programmer	1
	Networking	0.9
	Multimedia	0.8
	Bidang lain	0.4
	Tidak ada	0.1

2.2. Bobot Kriteria

Kriteria yang digunakan pada penelitian ini terdiri dari, 1) Test tertulis, mahasiswa akan diuji logia berpikit, 2) Test wawancara, melihat kepercayaan diri mahasiswa, 3) IPK, memiliki indeks prestasi yang baik, 4) Keaktifan dikelas, untuk melihat cara mahasiswa beradaptasi, 5) Kepemimpinan, kemampuan mahasiswa dalam mengatur tim anggotanya, 6) UKM (Unit Kegiatan Mahasiswa), keterlibatan dalam kegiatan mahasiswa, 7) Sertifikat Keahlian, aktif melakukan pengembangan diri. Dimana bobot kriteria masing – masing sebagai berikut, 20%, 20%, 20%, 10%, 10% dan 10%.

2.3. Proses Clustering

Berdasarkan skenario diatas, langkah pertama yang dilakukan adalah proses pengelompokan dengan metode *K-Means*. Tujuannya adalah untuk mengidentifikasi kelompok objek dengan karakteristik yang mirip dan berbeda dengan objek lainnya, sehingga setiap cluster memiliki sifat homogen [18]. *K-Means cluster* akan membagi sejumlah n objek kedalam kelompok k yang pengamatannya berdasarkan nilai rata – rata (mean) [19]. Nilai k dapat ditentukan sendiri ataupun dapat menggunakan pendekatan *rule-of-thumbs* [13] dalam penelitian ini menggunakan nilai $k = 2$ dimana kluster pertama untuk objek dengan karakteristik yang memenuhi syarat dan sebaliknya. Untuk *centroid* awal menggunakan nilai Maksimum dan Minimum masing – masing untuk kluster yang memenuhi persyaratan dan tidak. Algoritme *K-Means* dapat diuraikan berikut ini:

1. Menentukan jumlah kluster yang akan digunakan dalam hal ini $k = 2$, yang terdiri dari kluster yang memenuhi syarat dan sebaliknya.
2. Menentukan nilai *centroid* awal dengan menggunakan nilai maksimum dan minimum dari seluruh kriteria yang ada.
3. Hitung nilai jarak dari data ke kedua centroid awal tadi menggunakan *euclidian distance*.
4. Kemudian tentukan nilai terdekat dengan centroid tersebut sehingga dapat dikelompokan dalam kedua cluster.
5. Lakukan proses ini sampai nilai konvergen dimana tidak ada lagi *cluster* yang berpindah.

2.4. Proses Perangkingan

Setelah proses clustering selesai maka anggota yang ada pada kluster pertama akan dirangking dengan menggunakan metode TOPSIS. Metode ***Technique for Order Preferences by Similarity to Ideal Solution*** (TOPSIS) diusulkan pertama kali oleh Hwang dan Yoon [20] dan disempurnakan oleh Chen dan Hwang [21]. Prinsip dasar metode TOPSIS adalah menentukan alternatif dengan jarak terdekat dari solusi ideal positif dan jarak terjauh dari solusi ideal negatif. Tiga varian dari metode TOPSIS, 1) model konvensional atau C-TOPSIS oleh Hwang dan Yoon, 2) A-TOPSIS, metode yang disesuaikan dan 3) Metode yang dimodifikasi atau M-TOPSIS [21]. Penelitian ini menggunakan C-TOPSIS yang diusulkan oleh Hwang dan Yoon. Berikut algoritma dari TOPSIS [3], [22]. Adapun proses untuk melakukan perangkingan dengan TOPSIS adalah sebagai berikut,

Langkah 1: Membentuk normalisasi matriks keputusan. Penelitian ini menggunakan *Vector Normalization* seperti terlihat pada persamaan 1 dan 2 [23],

Kriteria benefit,

$$n_{ij} = \frac{r_{ij}}{\sqrt{\sum_{i=1}^m r_{ij}^2}}, i = 1, \dots, m; j = 1, \dots, n \quad (1)$$

Kriteria Cost,

$$n_{ij} = \frac{r_{ij}}{\sqrt{\sum_{i=1}^m r_{ij}^2}}, i = 1, \dots, m; j = 1, \dots, n \quad (2)$$

Langkah 2: Membuat matriks keputusan bobot ternormalisasi, dengan cara mengalikan nilai normalisasi dengan bobot kriteria,

$$v_{ij} = w_j \cdot r_{ij} \quad (3)$$

dimana, w_j bobot untuk kriteria j .

Langkah 3: Tentukan nilai solusi ideal positif dan ideal negatif.

Untuk Solusi Ideal Positif,

$$A^* = \{v_1^*, \dots, v_n^*\}, \quad (4)$$

dimana, $v_j^* = \{\max(v_{ij}) \text{ if } j \in J; \min(v_{ij}) \text{ if } j \in J'\}$

Untuk Solusi Ideal Negatif,

$$A' = \{v'_1, \dots, v'_n\}, \quad (5)$$

dimana, $v'_j = \{\min(v_{ij}) \text{ if } j \in J; \max(v_{ij}) \text{ if } j \in J'\}$

Langkah 4: Hitung *separation measure* untuk setiap alternatif, dimana pada bagian ini selain menggunakan *Eucliden Distance* (default dari TOPSIS) juga menggunakan *Manhattan Distance*.

Untuk alternatif Ideal positif adalah,

$$S_i^* = \sqrt{\sum (v_j^* - v_{ij})^2}, i = 1, \dots, m \quad (\text{Eucliden distance}) \quad (6)$$

$$S_i^* = \sum |v_j^* - v_{ij}|, i = 1, \dots, m \quad (\text{Manhattan distance}) \quad (7)$$

Untuk alternatif Ideal negatif adalah,

$$S_i' = \sqrt{\sum (v'_j - v_{ij})^2}, i = 1, \dots, m \quad (\text{Eucliden distance}) \quad (8)$$

$$S_i' = \sum |v'_j - v_{ij}|, i = 1, \dots, m \quad (\text{Manhattan distance}) \quad (9)$$

Langkah 5: Hitung nilai relatif *Closeness Coefficient* untuk solusi ideal C_i^*

$$C_i^* = \frac{S_i'}{(S_i^* + S_i')} \quad (10)$$

dimana, $0 < C_i^* < 1$

4. Hasil dan Pembahasan

4.1. Hasil Perhitungan

Berdasarkan skenario yang diusulkan maka proses pertama yang akan dilakukan adalah pemisahan objek berdasarkan kemiripan dari karakteristik setiap objek dalam hal ini proses *cluster*. *K-Means* digunakan untuk proses *clustering* dengan menggunakan nilai k adalah 2. Untuk *centroid* awal digunakan nilai Maksimum untuk group yang memenuhi syarat atau Layak dan nilai Minimum untuk group yang tidak memenuhi syarat. Tabel 2, merupakan data awal mahasiswa hasil kuisioner yang dibuat, Tabel 3 merupakan hasil konversi ke nilai numerik sesuai nilai pada Tabel 1.

Tabel 2. Data Awal dari Mahasiswa

NO.	NIM	TES TERTULIS	TES WAWANCARA	IPK	KEAKTIFAN DIKELAS	KEPEMIMPINAN	UKM	SERTIFIKAT KEAHLIAN
1.	1705001	10	Tidak	3.49	Kurang	Tidak Ada	Tidak	Bidang Lain
2.	1705002	50	Cukup	3.82	Sedikit	Seksi / Bagian	Kurang	Bidang Lain
3.	1705004	33	Cukup	3.71	Aktif	Seksi / Bagian	Tidak	Networking
4.	1705005	46	Kurang	3.54	Kurang	Tidak Ada	Tidak	Tidak Ada
5.	1705006	48	Kurang	3.69	Kurang	Tidak Ada	Kurang	Bidang Lain
...
49.	1705075	0	Tidak	3.43	Kurang	Tidak Ada	Tidak	Tidak Ada
50.	1705076	0	Tidak	3.27	Kurang	Tidak Ada	Tidak	Tidak Ada
51.	1705077	38	Kurang	3.82	Sedikit	Seksi / Bagian	Kurang	Bidang Lain
52.	1705079	25	Kurang	3.29	Kurang	Tidak Ada	Tidak	Tidak Ada
53.	1705080	35	Tidak	3.65	Kurang	Sekretaris	Tidak	Bidang Lain
54.	1705082	71	Memuaskan	3.95	Aktif	Bendahara	Aktif	Bidang Lain
55.	1705083	5	Tidak	3.10	Kurang	Tidak Ada	Tidak	Bidang Lain

Tabel 3. Matriks penilaian setelah di konversi ke nilai Numerik

NIM	TT	TW	IPK	KD	KP	UKM	SK
1705001	10	0.4	3.49	0.2	0.1	0.2	0.4
1705002	50	0.8	3.82	0.9	0.4	0.6	0.4
1705004	33	0.8	3.71	1.0	0.4	0.2	0.9
1705005	46	0.6	3.54	0.2	0.1	0.2	0.1
1705006	48	0.6	3.69	0.2	0.1	0.6	0.4
...
1705077	38	0.6	3.82	0.9	0.4	0.6	0.4
1705079	25	0.6	3.29	0.2	0.1	0.2	0.1
1705080	35	0.4	3.65	0.2	0.8	0.2	0.4
1705082	71	1.0	3.95	1.0	0.6	1.0	0.4
1705083	5.0	0.4	3.10	0.2	0.1	0.2	0.4

Ket: TT= Tes Tertulis, TW=Tes Wawancara, KD=Keaktifan dikelas, SK=Sertifikat keahlian.

Setelah data dikonversi ke dalam bentuk numerik, maka tentukan nilai *centroid* awal C1 dan C2 dengan mengambil nilai Maksimum dan Minimum dari setiap kriteria seperti pada Table 4,

Tabel 4. Centroid awal berdasarkan nilai Maksimum dan Minimum

		TT	TW	IPK	KD	KP	UKM	SK
C1	MAKS	75.00	1.00	3.95	1.00	1.00	1.00	1.00
C2	MIN	2.00	0.40	2.90	0.20	0.10	0.20	0.10

Kemudian hitung jarak dengan *Euclidian distance*. Selanjutnya periksa nilai jarak yang terdekat dengan setiap *centroid* C1 dan C2 untuk dimasukkan kedalam *cluster* K1 dan K2. Table 5, merupakan perhitungan jarak dan kluster. Kemudian *update centroid* dengan mengambil nilai rata – rata dari setiap anggota yang ada pada kluster 1 dan 2. Untuk *centroid* 1 akan dihitung rata – rata dari *cluster* 1 dan *centroid* 2 dihitung nilai rata – rata dari *cluster* 2 sehingga akan diperoleh nilai C1 dan C2 yang baru. Kemudian lakukan proses perhitungan jarak sampai nilai mencapai konvergensi dimana elemen dalam setiap *cluster* tidak berpindah.

Tabel 5. Perhitungan Jarak dan cluster untuk iterasi pertama

Dist.			Dist.			Dist.		
d1	d2	Clust.	d1	d2	Clust.	d1	d2	Clust.
65.023	8.027	2	24.012	49.029	1	27.028	46.013	1
25.019	48.019	1	73.022	0.920	2	73.025	0.450	2
42.013	31.035	2	60.012	13.077	2	73.025	0.390	2
29.056	44.005	1	65.021	8.050	2	25.002	48.035	1
27.041	46.010	1	47.012	26.026	2	73.026	0.320	2
72.011	1.488	2	37.028	36.019	2	70.013	3.201	2
73.030	0.000	2	55.021	18.019	2	45.026	28.012	2
35.032	38.012	1	22.024	51.017	1	10.019	63.027	1
30.023	43.018	1	73.025	0.470	2	73.024	0.570	2
73.023	0.790	2	70.013	3.213	2	73.019	1.090	2
10.098	63.011	1	70.020	3.079	2	55.009	18.050	2
65.022	8.057	2	73.026	0.330	2	73.024	0.530	2
0.304	73.024	1	67.013	6.087	2	73.025	0.370	2
39.035	34.014	2	47.010	26.035	2	37.014	36.024	2
60.013	13.051	2	35.022	38.028	1	50.035	23.004	2
5.057	68.020	1	73.013	1.178	2	40.027	33.017	2
73.024	0.530	2	60.109	13.028	2	4.064	69.022	1
45.018	28.025	2	73.027	0.220	2	70.025	3.022	2
35.019	38.027	1						

Ket: d1=Jarak dari data ke Centroid 1; d2=Jarak dari data ke *Centroid* 2.

Setelah melakukan iterasi sebanyak enam kali maka konvergensi terjadi pada iterasi yang ke 5 dan nilai kluster sudah tidak berubah. Sehingga anggota kluster 1 terdiri dari 25 responden yang akan dirangking sedangkan sisanya masuk kedalam kluster 2.

Selanjutnya, proses perangkingan menggunakan metode TOPSIS. Langkah pertama adalah Normalisasi menggunakan normalisasi *Vector*. Normalisasi dilakukan terhadap data yang masuk dalam kluster pertama yaitu sejumlah 25 responden menggunakan persamaan 1, kemudian hitung matriks bobot ternormalisasi yang merupakan perkalian antara nilai yang ada pada kriteria dengan bobot masing – masing kriteria seperti pada persamaan 3. Kemudian tentukan nilai ideal positif dan negatif seperti pada persamaan 4 dan 5. Selanjutnya hitung nilai S^* dan S' , menggunakan jarak *Euclidean* (Persamaan 6 dan 8) dan *Manhattan* (Persamaan 7 dan 9). Terakhir menghitung nilai C^* untuk jarak *Euclidean* dan *Manhattan distance*. Hasil akhir dari kedua perhitungan tersebut terlihat pada Table 6, yang terdiri dari nilai S^* , S' , C^* serta nilai perangkingan akhir.

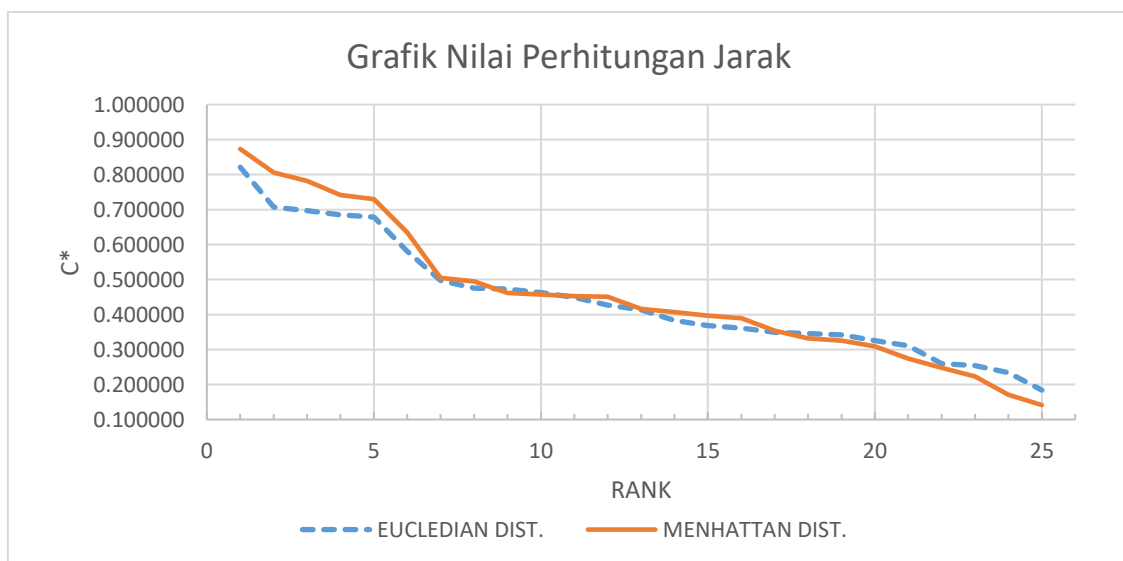
Tabel 6. Nilai S^* , S' dan C^* dari *Euclidean* dan *Manhattan Distance*

Respon- dent	Euclidean Distance			Rank	Manhattan Distance			Rank
	S^*	S'	C^*		S^*	S'	C^*	
1705002	0.041	0.038	0.476	8	0.092	0.094	0.505	7
1705004	0.050	0.041	0.449	11	0.100	0.086	0.461	9
1705005	0.064	0.020	0.234	24	0.154	0.032	0.171	24
1705006	0.055	0.027	0.325	20	0.128	0.057	0.309	20
1705009	0.061	0.021	0.260	22	0.140	0.046	0.248	22
1705010	0.044	0.038	0.463	10	0.102	0.084	0.452	11
1705014	0.046	0.046	0.498	7	0.094	0.092	0.495	8
1705018	0.014	0.066	0.821	1	0.024	0.162	0.873	1
1705020	0.069	0.016	0.184	25	0.160	0.026	0.141	25
1705022	0.027	0.059	0.685	4	0.050	0.136	0.729	5
1705024	0.056	0.031	0.361	16	0.120	0.066	0.354	17

Respondent	Euclidean Distance			Rank	Manhattan Distance			Rank
	S*	S'	C*		S*	S'	C*	
1705025	0.056	0.033	0.368	15	0.112	0.074	0.396	15
1705026	0.035	0.049	0.581	6	0.068	0.118	0.635	6
1705037	0.057	0.026	0.311	21	0.124	0.062	0.332	18
1705038	0.058	0.030	0.342	19	0.125	0.060	0.325	19
1705040	0.045	0.034	0.427	12	0.101	0.085	0.457	10
1705050	0.053	0.038	0.383	14	0.110	0.076	0.407	14
1705052	0.054	0.038	0.413	13	0.108	0.077	0.416	13
1705057	0.043	0.038	0.473	9	0.102	0.084	0.451	12
1705062	0.026	0.055	0.678	5	0.048	0.138	0.742	4
1705067	0.060	0.021	0.254	23	0.144	0.041	0.223	23
1705068	0.026	0.062	0.707	2	0.036	0.150	0.806	2
1705077	0.051	0.028	0.350	17	0.113	0.072	0.389	16
1705080	0.059	0.031	0.345	18	0.135	0.051	0.274	21
1705082	0.027	0.061	0.697	3	0.040	0.145	0.782	3

4.2. Pembahasan

Dalam penelitian ini, proses *cluster* digunakan untuk mengeliminasi responden yang tidak memenuhi persyaratan yang ditentukan. Dengan menggunakan *K-Means*, maka terdapat 25 anggota pada *cluster* 1 yang memenuhi persyaratan sedangkan *cluster* 2 berjumlah 30. Jumlah mahasiswa yang akan diberangkatkan hanya terdiri dari 20 orang, sehingga seluruh anggota pada *cluster* 1 akan di rangking untuk menentukan peringkat mahasiswa yang dapat mengikuti kegiatan tersebut. Terdapat beberapa hal penting yang akan didiskusikan yaitu: 1) jika dibandingkan dengan kenyataan dimana mahasiswa yang berangkat dari *cluster* 1 sebanyak 16 orang dari 25 anggota dan 4 orang diambil dari *cluster* 2 berdasarkan wawancara akhir beserta dengan penilaian dari pimpinan. 2) dalam proses perankingan dengan menggunakan metode MCDM yaitu TOPSIS dapat dihasilkan rangking yang sesuai namun dengan perbedaan metode pengukuran menghasilkan rangking yang berbeda, dimana *Manhattan distance* memberikan penilaian yang lebih besar sekitar 60% dari *Euclidian*. Hasil tersebut sejalan dengan penelitian [24]. Gambar 1 menunjukkan perbedaan nilai dari *Euclidian* dan *Manhattan distance*. Secara keseluruhan *Manhattan distance* dapat digunakan untuk menggantikan pengukuran ideal positif dan negatif dari *Euclidian distance*.



Gambar 1. Grafik Perbandingan Nilai *Euclidian Distance* dan *Manhattan Distance*

3) penggunaan metode jarak untuk mengukur nilai similarity pada metode TOPSIS memberikan pengaruh kepada hasil perangkingan dimana terdapat 5 anggota dari *cluster* 1 yang berada diluar ranking sehingga tidak dapat diikutsertakan, dimana dalam realisasi terdapat 1 anggota dengan nim 1705080 yang dengan menggunakan jarak *Euclidian* menduduki ranking ke 21 sedangkan pada *Manhattan distance* menduduki ranking ke 18. Oleh karena itu jika menggunakan *euclidian distance* maka mahasiswa ini tidak dapat ikut dalam kegiatan tersebut, sedangkan jika menggunakan *Manhattan distance* mahasiswa ini terpilih dalam urutan ke 18. Hal ini membuktikan bahwa penentuan jarak pada perhitungan metode TOPSIS sangat berpengaruh.

5. Simpulan

Dari hasil penelitian dan diskusi yang dilakukan maka dapat disimpulkan bahwa penggabungan metode *cluster* dengan metode MCDM dapat menjadi standar perhitungan untuk pemilihan dan perangkingan suatu kasus tertentu. Akan tetapi karena sistem pendukung keputusan masih memberikan kewenangan besar bagi pimpinan dalam menentukan alternatif yang akan dipilih maka pengambilan data dari luar klaster yang telah terbentuk masih dapat dilakukan. Dengan menggunakan metode TOPSIS dapat dilakukan proses perangkingan dengan lebih akurat dan objektif selain itu penggunaan *Manhattan distance* memberikan nilai yang lebih tinggi dari pada *Euclidian distance* sekitar 60% dari total data klaster satu. Setelah proses perangkingan dengan *Euclidian distance* terdapat 70% dari 20 mahasiswa yang terpilih dan 30% diambil dari cluster pertama dan kedua. Dengan *Manhattan distance* terdapat 75% dari 20 mahasiswa yang terpilih dan 25% diambil dari klaster pertama dan kedua. Dengan demikian *Manhattan distance* memberikan hasil yang mendekati dengan realisasi mahasiswa yang berangkat mengikuti kegiatan kunjungan industri. Dengan menggunakan kombinasi kedua metode diharapkan bagi pimpinan untuk dapat menggunakan anggota yang masuk dalam klaster pertama dan tidak perlu mengambil data dari klaster lainnya.

Daftar Referensi

- [1] L. A. Manurung, "Kunjungan Industri Program Vokasi - Learning Experience.," 2020. <https://kalbelearningcentre.kalbe.co.id/News/ArtMID/547/ArticleID/30/Kunjungan-Industri-Program-Vokasi> (accessed Feb. 21, 2021).
- [2] V. Nastiti1, A. C. Sukartiko, and N. E. Kristanti, "Analisis Kepuasan Pengunjung Terhadap Pelayanan Kunjungan Industri Di PT. Sido Muncul," *Repository, Universitas Gadjah Mada*, 2016.
- [3] A. B. Hastuti, E. Utami, and E. T. Luthfi, "Implementasi Metode Fuzzy C-Means Dan Topsis Dalam Membangun Sistem Pendukung Keputusan Penentuan Jurusan SMA (Studi Kasus : Penentuan Jurusan Di SMA Negeri 1 Wonosari)," *DASI*, vol. 14, no. 2, pp. 9-15, 2013.
- [4] N. D. Budiana, R. R. A. Siregar, and M. N. I. Susanti, "Penetapan Instruktur Diklat Menggunakan Metode Clustering K-Means dan Topsis Pada PT PLN (Persero) Udiklat Jakarta," *petir*, vol. 12, no. 2, pp. 111–121, Aug. 2019, doi: 10.33322/petir.v12i2.454.
- [5] Mirfan, "Sistem Pendukung Keputusan Pemilihan Destinasi Wisata Berbasis Web Dengan Algoritma K-Means Clustering Dan TOPSIS," *Jurnal INSTEK*, vol. 5, no. 2, pp. 240–250, Oktober 2020.
- [6] E. Daniati and H. Utama, "Clustering K Means for Criteria Weighting With Improvement Result of Alternative Decisions Using SAW and TOPSIS," in *2019 4th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, Yogyakarta, Indonesia, pp. 73–78, 2019, doi: 10.1109/ICITISEE48480.2019.9003858.
- [7] K. Khomsatun, D. Ikhsan, M. Ali, and K. Kursini, "Sistem Pengambilan Keputusan Pemilihan Lahan Tanam Di Kabupaten Wonosobo Dengan K-Means Clustering Dan TOPSIS," *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, vol. 9, no. 1, pp. 55-62, Apr. 2020, doi: 10.23887/janapati.v9i1.23073.
- [8] A. Masruro and E. T. Luthfi, "Sistem Penunjang Keputusan Penentuan Lokasi Wisata Menggunakan K-Means Clustering Dan TOPSIS," vol. 15, no. 04, pp. 1-5, Desember 2011.
- [9] M. A. Raharja, I. K. A. Surya, and I. K. A. Mogi, "Clustering Customer For Determine Market Strategy Using K-Means And TOPSIS: Case Study," *Vol. 2*, pp. 61-71, 2022.

- [10] F. Sun and J. Yu, "Improved energy performance evaluating and ranking approach for office buildings using Simple-normalization, Entropy-based TOPSIS and K-means method," *Energy Reports*, vol. 7, pp. 1560–1570, Nov. 2021, doi: 10.1016/j.egy.2021.03.007.
- [11] B. Trstenjak, A. Rijana K. Kavran, and I. Bujan, "Evaluation of Croatian Development Strategies Using SWOT Analyses with Fuzzy TOPSIS Method and K-Means Methods," *JOEBM*, vol. 3, no. 7, pp. 687–693, 2015, doi: 10.7763/JOEBM.2015.V3.267.
- [12] Havaluddin *et al.*, "A Performance Comparison of Euclidean, Manhattan and Minkowski Distances in K-Means Clustering," in *2020 6th International Conference on Science in Information Technology (ICSITech)*, Palu, Indonesia, pp. 184–188, 2020. doi: 10.1109/ICSITech49800.2020.9392053.
- [13] M. Nishom, "Perbandingan Akurasi Euclidean Distance, Minkowski Distance, dan Manhattan Distance pada Algoritma K-Means Clustering berbasis Chi-Square," *jpit*, vol. 4, no. 1, pp. 20–24, Jan. 2019, doi: 10.30591/jpit.v4i1.1253.
- [14] A. Setiawan, "Perbandingan Penggunaan Jarak Manhattan, Jarak Euclid, dan Jarak Minkowski dalam Klasifikasi Menggunakan Metode KNN pada Data Iris," *juses*, vol. 5, no. 1, pp. 28–37, May 2022, doi: 10.24246/juses.v5i1p28-37.
- [15] W.-Y. Chiu, G. G. Yen, and T.-K. Juan, "Minimum Manhattan Distance Approach to Multiple Criteria Decision Making in Multiobjective Optimization Problems," *IEEE Trans. Evol. Computat.*, vol. 20, no. 6, pp. 972–985, Dec. 2016, doi: 10.1109/TEVC.2016.2564158.
- [16] S. B. Sakur, M. Silangen, and E. H. Israel, "Penggunaan Metode Technique For Order Performance Of Similarity To Ideal Solution (TOPSIS) Dan Vector Normalization Pada Pemilihan Mahasiswa Kunjungan Industri," Politeknik Negeri Nusa Utara, Tahuna, Laporan Penelitian Unggulan Perguruan Tinggi 461/Sistem Informasi, Nov. 2021.
- [17] S. B. Sakur, "Data Excel Proses Analisis Kunjungan Industri Metode TOPSIS," *Repositori Stendy B. Sakur*, Nov. 21, 2021. <https://drive.google.com/file/d/1Z8UmqyJ7v5Nc4Y42zLWOv7kxjz0l78ny/view?usp=sharing> (accessed Nov. 21, 2021).
- [18] N. Arief, I. S. Damanik, and E. Irawan, "Penerapan Algoritma K-Medoids Dalam Mengelompokkan Tingkat Kasus Kejahatan di Setiap Provinsi," *KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 2, no. 3, pp. 111–116, 2021.
- [19] N. Islam, "Defining Homogenous Climate zones of Bangladesh using Cluster Analysis," vol. 6, no. 1, pp. 119–129, Feb. 2019.
- [20] C.-L. Hwang and K. Yoon, "Methods for multiple attribute decision making," in *Multiple attribute decision making*, Springer, pp. 58–191, 1981.
- [21] M. Abedi and G.-H. Norouzi, "A general framework of TOPSIS method for integration of airborne geophysics, satellite imagery, geochemical and geological data," *International Journal of Applied Earth Observation and Geoinformation*, vol. 46, pp. 31–44, Apr. 2016, doi: 10.1016/j.jag.2015.11.016.
- [22] M. Behzadian, S. Khanmohammadi Otaghsara, M. Yazdani, and J. Ignatius, "A state-of-the-art survey of TOPSIS applications," *Expert Systems with Applications*, vol. 39, no. 17, pp. 13051–13069, Dec. 2012, doi: 10.1016/j.eswa.2012.05.056.
- [23] D. M. Pavlicic, "Normalisation Affects the Results Of Madm Methods," vol. 11, no. 2, pp. 261–265, 2001.
- [24] D. Sinwar, R. Kaushik, "Study of Euclidean and Manhattan distance metrics using simple k-means clustering". *Int. J. Res. Appl. Sci. Eng. Technol*, vol. 5, no. 2, pp. 270-274, 2014.