

## Implementasi Metode K-Means untuk Pengelompokan Rekomendasi Tugas Akhir

Haviluddin <sup>1,\*</sup>), Suryani Junita Patandianan <sup>2)</sup>, Gubtha Mahendra Putra <sup>3)</sup>, Novianti Puspitasari <sup>4)</sup>, Herman Santoso Pakpahan <sup>5)</sup>

Program Studi Informatika, Fakultas Teknik, Universitas Mulawarman  
Kampus Universitas Mulawarman, Gunung Kelua, Jalan Sambaliung No.9 Samarinda 75119  
e-mail : haviluddin@unmul.ac.id <sup>1)</sup>; suryani.patandianan@gmail.com <sup>2)</sup>; gubthamp@fkti.unmul.ac.id <sup>3)</sup>;  
novia.ftik.unmul@gmail.com <sup>4)</sup>; pakpahan.herman891@gmail.com <sup>5)</sup>

### ABSTRAK

Tugas akhir (TA) merupakan salah satu syarat dalam penyelesaian jenjang pendidikan formal di suatu perguruan tinggi. Namun, kebanyakan mahasiswa mengalami kesulitan dalam menentukan area penelitian sehingga TA yang diambil kurang atau tidak sesuai dengan kemampuan mereka. Penelitian ini bertujuan untuk merekomendasikan area penelitian TA bagi mahasiswa berdasarkan data nilai A, B dan C pada 10 Mata Kuliah Wajib (MKW) yang diperoleh selama 6 semester untuk angkatan 2015, 2016, 2017 dan 2018 pada mahasiswa Program Studi Informatika, Fakultas Teknik, Universitas Mulawarman. Analisa pengelompokan menggunakan metode *K-Means*. Berdasarkan percobaan, pengelompokan area penelitian TA dilakukan dengan 3 cluster (C), yaitu C1 adalah sedikit, beranggotakan 1 MKW; C2 adalah sedang, beranggotakan 6 MKW; dan C3 adalah banyak, beranggotakan 3 MKW telah diperoleh. Pengujian akurasi *cluster* menggunakan metode *Sum of Squared Errors* (SSE) sebesar 0.6566 dan metode *Silhouette Coefficient* (SC) sebesar 5.8329 telah didapatkan. Hal ini menunjukkan bahwa nilai MKW juga memiliki pengaruh dalam menentukan TA.

Kata Kunci – Rekomendasi, Tugas akhir (TA), *K-Means*, *Sum Squared Error* (SSE), *Silhouette Coefficient* (SC)

### 1. PENDAHULUAN

Setiap mahasiswa yang menempuh jenjang perguruan tinggi, baik negeri maupun swasta, diwajibkan untuk menyelesaikan tugas akhir (TA) atau skripsi menjelang akhir masa studi sebagai prasyarat kelulusan mahasiswa. Namun demikian, mencari dan memahami materi skripsi yang akan dikerjakan menjadi salah satu hambatan. TA yang diambil seharusnya dapat disesuaikan dengan kemampuan mahasiswa dalam memahami suatu mata kuliah yang pernah ditempuh (Farokhah & Aditya, 2017; Muttaqin & Defriani, 2020). Namun, sebagian besar mahasiswa mengambil TA karena mengikuti rekan seangkatan yang sudah lebih dulu mendapatkan judul atau referensi judul dari internet yang tidak sesuai dengan kemampuan mereka. Hal ini dapat menghambat mahasiswa tersebut dalam proses pengerjaan dan penyelesaian TA (Siswanto & Sampurna, 2013). Lebih lanjut, nilai yang diperoleh pada mata kuliah wajib (MKW) dengan TA memiliki hubungan yang erat. Dimana, nilai yang diperoleh mengindikasikan kemampuan mahasiswa dalam memahami mata kuliah tersebut. Oleh karena itu, analisa keterhubungan nilai MKW yang pernah ditempuh selama kuliah dapat membantu dalam menentukan TA sangat diperlukan.

Berbagai algoritma cerdas untuk menganalisa telah banyak diimplementasikan. Salah satunya adalah algoritma *K-Means* yang merupakan algoritma pengelompokan *iterative* yang sederhana diimplementasikan, relatif cepat, dan mudah beradaptasi (Parlina, Windarto, Wanto, & Lubis, 2018). Keterhubungan nilai MKW dan TA sangat perlu dianalisa agar mahasiswa, dosen dan program studi dapat mengetahui kondisi setiap mahasiswa. Berbagai penelitian yang menggunakan algoritma *K-*

*Means* telah dilakukan oleh peneliti Hermawan, Ugiarto, & Puspitasari (Hermawan, Ugiarto, & Puspitasari, 2017) telah menggunakan algoritma *K-Means* untuk mengelompokkan kinerja 32 asisten laboratorium menjadi 3 (tiga) *cluster* yaitu C1 = sangat bagus, C2 = bagus dan C3 = kurang bagus. Hasil penelitian memperlihatkan bahwa 12 asisten (C1), 14 asisten (C2), dan 6 asisten (C3). Peneliti Kurnia Bakti & Indriyatno (Kurnia Bakti & Indriyatno, 2017) telah menggunakan algoritma *K-Means* untuk menganalisa dokumen tugas akhir dari 4 (empat) program studi Diploma tiga (D3) yaitu D3 Teknik Komputer, D3 Farmasi, D3 Akuntansi dan D3 Kebidanan, Politeknik Harapan Bersama. Data diperoleh secara acak dari tahun 2013-2015 dengan masing-masing program studi sebanyak 50 laporan TA. Berdasarkan percobaan indikator jarak antar klaster yang dihasilkan sebesar 0.001 dengan metode *Davies Bouldin Index*. Hasil penelitian memperlihatkan bahwa terdapat 4 (empat) *cluster* yang digunakan yaitu C0 = 46 *items*, C1 = 66 *items*, C2 = 45 *items*, dan C3 = 43 *items*. Peneliti Asroni (Asroni, 2015) telah menggunakan algoritma *K-Means* untuk mengelompokkan mahasiswa berdasarkan nilai akademik pada jurusan Teknik Informatika UMM Magelang. Data penelitian yang digunakan adalah 124 jumlah *instance* yang memiliki 5 atribut yaitu NIM, nilai MK algoritma dan pemrograman, MK fisika dasar, MK kalkulus 1, dan IPK. Hasil penelitian memperlihatkan bahwa *cluster* 1 adalah IPK tertinggi dapat digunakan untuk memilih 5 mahasiswa untuk mewakili lomba. Peneliti Muzakir (Muzakir, 2014) telah menggunakan algoritma *K-Means* untuk menganalisa penentuan beasiswa bagi 22 mahasiswa dan penghasilan orang tua di Kabupaten Musi Banyuasin. Berdasarkan hasil

penelitian menunjukkan bahwa terdapat 3 (tiga) *cluster* yang terbentuk yaitu C1 = sangat baik, C2 = baik, dan C3 = kurang baik. Hasil penelitian memperlihatkan bahwa diperoleh sebanyak 2 siswa berada pada C2 dan 20 siswa berada pada C3. Peneliti Adrianto (Adrianto, 2016) telah menggunakan algoritma *K-Means* untuk menganalisa pemilihan jalur peminatan sesuai kemampuan pada Program Studi Teknik Informatika, Universitas Dian Nuswantoro. Berdasarkan hasil percobaan merekomendasikan sebanyak 5 (lima) jalur peminatan terdiri dari C0 = RPLD, C1 = RPLD, C2 = SC, C3 = SC dan C4 = SC.

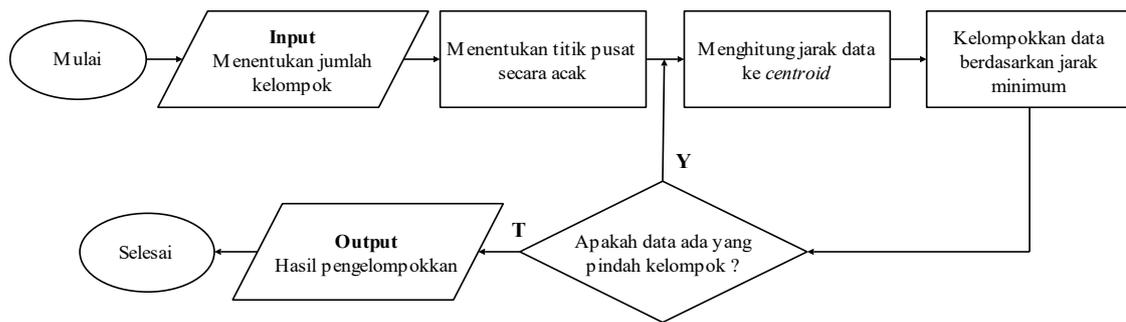
Artikel ini menggunakan algoritma *K-Means* untuk menganalisa keterhubungan nilai MKW dengan TA mahasiswa. Tujuan penelitian ini adalah untuk membantu mahasiswa, dosen dan program studi dalam memberikan masukan dalam merencanakan

penelitian sehingga diharapkan TA yang ditempuh dapat dikerjakan dengan baik, tepat waktu dan sesuai dengan kemampuan mahasiswa. Artikel ini terdiri dari motivasi penulisan artikel pada bagian pertama. Kedua, menjelaskan model kerja algoritma *K-Means*. Ketiga, analisa hasil percobaan. Kesimpulan penelitian pada bagian akhir.

## 2. TINJAUAN PUSAKA

### A. Algoritma *K-Means*

Algoritma *K-Means* merupakan salah satu algoritma pengelompokan (*clustering*) berbasis metode *non-hierarchy* yang mempartisi data dan membentuk satu atau lebih kelompok yang memiliki kesamaan (Hasanah, Ugiarto, & Puspitasari, 2017; Purnawansyah, Haviluddin, Gafar, & Tahyudin, 2017; Putra, Santosa, & Kusumawardani, 2015).



Gambar 1. Flowchart *K-Means Clustering*

Gambar 1, memperlihatkan alur analisa dalam penerapan rekomendasi TA menggunakan algoritma *K-Means*. Pertama, menentukan jumlah *cluster* berdasarkan nilai A, B dan C. Dalam penelitian ini, sebanyak 3 *cluster* yaitu C1 adalah sedikit, C2 adalah sedang, dan C3 adalah banyak. Kedua, menentukan *centroid* awal berdasarkan nilai maksimal, rata-rata, dan minimal yang diperoleh dari data. Ketiga, menghitung jarak data dengan *centroid*. Dalam penelitian ini, metode *Euclidean Distance* (ED) telah digunakan, Persamaan 1.

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \dots\dots\dots(1)$$

Dimana, *i* adalah objek; *x, y* adalah koordinat objek.

Keempat, mengelompokkan data ke *centroid* terdekat. Terakhir, menghitung kembali nilai *centroid* hingga data tidak mengalami berubah lagi.

### B. Perhitungan Akurasi

Pengujian model dilakukan untuk mengetahui seberapa dekat relasi antara objek dalam sebuah *cluster* dan seberapa jauh sebuah *cluster* terpisah dengan *cluster* lain. Dalam penelitian ini, uji *cluster* menggunakan *Sum of Squared Errors* (SSE) dan *Silhouette Coefficient* (SC). SSE digunakan untuk menguji jumlah *cluster* yang ideal (Kusuma, 2015; Purnawansyah & Haviluddin, 2017; Puspitasari &

Haviluddin, 2016; Sari & Devianto, 2014). Sedangkan, SC merupakan metode evaluasi untuk mengetahui kekuatan dan ketepatan suatu hasil *cluster* (Anggara, Sujiani, & Helfi, 2016).

#### 1) *Sum of Squared Errors (SSE)*

Dalam penelitian ini, metode SSE untuk pengukuran jarak akurasi untuk menghitung selisih total dari nilai sebenarnya. Nilai SSE yang mendekati 0 menandakan bahwa model tersebut mempunyai komponen kesalahan acak terkecil dan digunakan untuk peramalan terhadap suatu model yang diamati (Purnawansyah & Haviluddin, 2017; Puspitasari & Haviluddin, 2016; Sari & Devianto, 2014; Sugiyama, 2015). Perhitungan metode SSE menggunakan Persamaan 2.

$$SSE = \sum_{i=1}^n (X_i - Y_i)^2 \dots\dots\dots(2)$$

Dimana, *X* adalah nilai aktual yang sebenarnya; *Y* adalah nilai yang tercapai.

#### 2) *Silhouette Coefficient (SC)*

Dalam penelitian ini, evaluasi metode SC digunakan untuk menguji kualitas dari 3 dan 4 *cluster*. Dimana, nilai SC terbaik digunakan untuk mengukur kualitas *cluster* (Gentle, Kaufman, & Rousseuw, 1991; Sugiyama, 2015). Perhitungan metode SC menggunakan Persamaan 3.

$$S(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \dots\dots\dots(3)$$

Dimana,  $S(i)$  adalah hasil nilai SC;  $b(i)$  adalah nilai minimum dari rata-rata jarak;  $a(i)$  adalah perbedaan rata-rata objek ( $i$ ) ke semua objek lain pada  $a$ .

**C. Sampling Data**

Dalam penelitian ini, data nilai A, B dan C dari 10 MKW yang berjumlah 488 telah diperoleh dari transkrip mahasiswa lulusan tahun 2015, 2016, 2017 dan 2018 pada program studi Informatika, Fakultas Teknik, Universitas Mulawarman. 10 MKW tersebut adalah mata kuliah yang dianggap mendukung dalam penentuan TA. Data nilai A, B dan C pada 10 MKW dinormalisasi untuk menyamakan rentang nilai dengan Persamaan 4. Adapun, data nilai dapat dilihat pada Tabel 1.

$$X' = \frac{0,8 (X - b)}{(a - b)} + 0,1 \dots\dots\dots(4)$$

Dimana,  $X'$  adalah data baru;  $X$  adalah data awal;  $b$  adalah nilai minimum;  $a$  adalah nilai maksimum.

Tabel 1. Nilai A, B dan C pada 10 MKW

No.	Mata Kuliah	2015	2016	2017	2018
1	Basis Data	150	92	83	92
2	Sistem Operasi	152	102	110	107
3	Sistem Jaringan Komputer	160	104	105	112
4	Sistem Keamanan Komputer	137	95	103	90
5	Kecerdasan Buatan	162	98	109	104
6	Algoritma dan Pemrograman	150	96	93	90
7	Pemrograman Visual	70	38	72	82
8	Pemrograman Berbasis Objek	158	84	92	102
9	Pemrograman Web	160	79	77	82
10	Struktur Data	149	80	96	86

**3. HASIL DAN PEMBAHASAN**

Dalam percobaan ini, metode *K-Means* telah diterapkan. Sebanyak 488 dataset pada 10 MKW dengan nilai A, B dan C telah diimplementasikan. Analisis yang dilakukan adalah mengelompokkan TA yang bertujuan untuk melihat kelompok MKW yang dianggap mendukung dalam penentuan TA. Berdasarkan percobaan, jumlah *cluster* yang telah digunakan adalah 3 *cluster* dengan kriteria *cluster* pertama (C1) adalah sedikit, *cluster* kedua (C2) adalah sedang, dan *cluster* ketiga (C3) adalah banyak. Sedangkan, Nilai *centroid* awal telah mengambil nilai terendah (minimum) untuk (C1), nilai rata-rata untuk (C2) dan nilai tertinggi (maksimum) untuk (C3). Adapun, nilai *centroid* yang telah diterapkan dapat dilihat pada Tabel 2.

Tabel 2. Nilai *Centroid* Awal

	Nilai <i>Centroid</i> Awal				
C1	0.3065	0.1000	0.3194	0.3839	Minimal
C2	0.7890	0.4148	0.4613	0.4658	Rata-rata
C3	0.9000	0.5258	0.5645	0.5774	Maksimal

Kemudian, menghitung jarak data ke *centroid* menggunakan metode ED dengan Persamaan (1). Hasil perhitungan jarak data ke *centroid* dapat dilihat pada Tabel 3. Selanjutnya, mengalokasikan masing-masing data ke dalam *centroid* terdekat. Sebaran data pada iterasi pertama dapat dilihat pada Tabel 4.

Proses selanjutnya, nilai *centroid* baru yang telah diperoleh pada iterasi pertama telah dihitung kembali. Nilai *centroid* baru pada iterasi pertama dapat dilihat pada Tabel 5. Proses menghitung *centroid* sampai tiap data pada *cluster* tidak berubah lagi telah dikerjakan. Hasil *cluster* pada iterasi terakhir dapat dilihat pada Tabel 6.

Jarak Data ke *Centroid* pada Iterasi Pertama

No.	Mata Kuliah	C1	C2	C3
1	Basis Data	0.6300	0.0871	0.2429
2	Sistem Operasi	0.7324	0.1695	0.0733
3	Sistem Jaringan Komputer	0.7754	0.1986	0.0347
4	Sistem Keamanan Komputer	0.6039	0.0981	0.2271
5	Kecerdasan Buatan	0.7611	0.1746	0.0648
6	Algoritma dan Pemrograman	0.6538	0.0749	0.2021
7	Pemrograman Visual	0.0000	0.5991	0.7945
8	Pemrograman Berbasis Objek	0.6661	0.0998	0.1870
9	Pemrograman Web	0.6389	0.1758	0.3301
10	Struktur Data	0.5982	0.0773	0.2594

Tabel 3. Hasil *Cluster* pada Iterasi Pertama

No.	Mata Kuliah	C1	C2	C3
1	Basis Data		*	
2	Sistem Operasi			*
3	Sistem Jaringan Komputer			*
4	Sistem Keamanan Komputer		*	
5	Kecerdasan Buatan			*
6	Algoritma dan Pemrograman		*	
7	Pemrograman Visual	*		
8	Pemrograman Berbasis Objek		*	
9	Pemrograman Web		*	
10	Struktur Data		*	

Tabel 4. Nilai *Centroid* Baru pada Iterasi Kedua

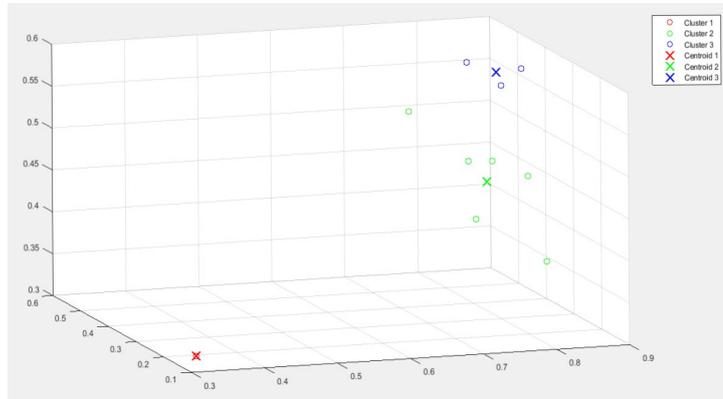
	Nilai <i>Centroid</i> Baru				
C1	0.3065	0.1000	0.3194	0.3839	Minimal
C2	0.7890	0.4148	0.4613	0.4658	Rata-rata
C3	0.9000	0.5258	0.5645	0.5774	Maksimal

Tabel 5. Hasil *Cluster* pada Iterasi Terakhir

No.	Mata Kuliah	C1	C2	C3
1	Basis Data		*	
2	Sistem Operasi			*
3	Sistem Jaringan Komputer			*
4	Sistem Keamanan Komputer		*	
5	Kecerdasan Buatan			*
6	Algoritma dan Pemrograman		*	
7	Pemrograman Visual	*		
8	Pemrograman Berbasis Objek		*	
9	Pemrograman Web		*	
10	Struktur Data		*	

Gambar 2, hasil analisa ditampilkan dalam bentuk grafik. Simbol “o” merupakan cluster dan simbol “x” merupakan nilai dari centroid. Terdapat 3 warna plot cluster yaitu merah (C1), hijau (C2), dan biru (C3). Selanjutnya, untuk mengetahui hasil cluster yang ideal dengan SSE telah dilakukan. Berdasarkan

percobaan, nilai SSE dengan 3 cluster sebesar 0.6566 % dan 4 cluster sebesar 1.1072 % telah diperoleh. Dalam percobaan ini menunjukkan bahwa 3 cluster memiliki nilai yang lebih baik dibandingkan dengan 4 cluster sehingga telah digunakan untuk analisa keterhubungan MKW dan TA.



Gambar 2. Sebaran C1, C2 dan C3

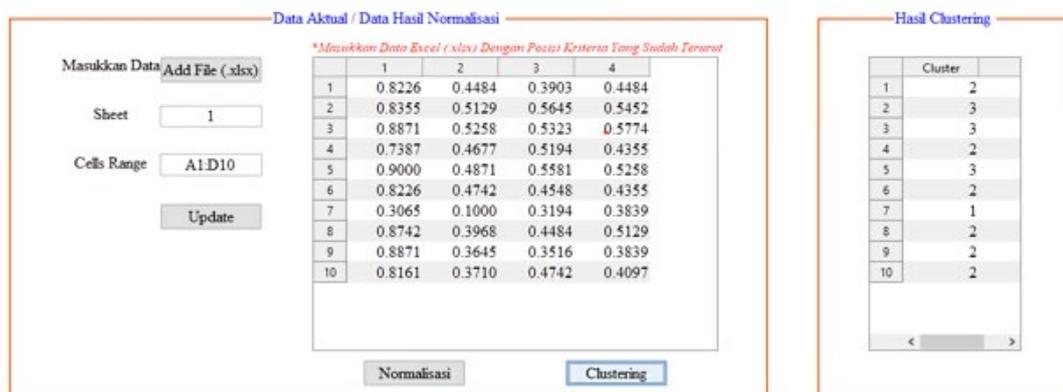
Dalam pengujian ini, keakuratan cluster juga telah diukur menggunakan metode SC. Berdasarkan hasil pengujian, nilai  $S_i$  Global 3 cluster adalah 0.5852% dan nilai  $S_i$  Global 4 cluster adalah 0.4591% telah diperoleh. Hal ini telah menunjukkan bahwa pengujian menggunakan 3 cluster telah memperoleh nilai yang lebih baik dibandingkan 4 cluster. Data nilai  $S_i$  pada 3 dan 4 cluster juga telah menunjukkan bahwa nilai  $S_i$  pada 3 cluster lebih kuat berdasarkan acuan dari kriteria subjektif pengukuran. Hal ini sesuai dengan pernyataan Kaufmann Roesseeuw (Gentle et al., 1991) yang telah menunjukkan bahwa hasil clustering dikatakan kuat jika nilai SC bernilai positif dan semakin mendekati nilai 1. Dengan kata lain, jika nilai SC yang baik terhadap suatu cluster maka hasilnya dapat digunakan sebagai dasar analisa. Hasil pengujian SC dapat dilihat pada Tabel 7.

2	0.138	0.187	0.260	0.071	0.189	0.178
9	7	1	3	3	9	
3	0.096	0.166	0.412	0	0.616	1
8	7	9	8			
4	0.118	0.161	0.267	0.073	0.179	0.132
2	4	7	0	0	3	
5	0.143	0.299	0.520	0.092	0.213	0.518
8	7	3	2	8	7	
6	0.105	0.221	0.525	0.074	0.183	0.515
2	8	5	0	5	6	
7	0.047	0.204	0.768	0.066	0.139	0.520
5	7	0	8	2	3	
8	0.046	0.222	0.789	0.083	0.187	0.555
9	9	6	2	2	6	
9	0.047	0.196	0.757	0.073	0.121	0.397
7	4	2	0	2	5	
10	0	0.631	1	0.081	0.132	0.386
		8		2	4	5
$S_i$						
Global	3 Cluster	0.585	4 Cluster	0.459		
al		2		1		

Tabel 6. Hasil Pengujian Silhouette Coefficient (SC)

Data ke-	3 Cluster			4 Cluster		
	$a_i$	$b_i$	$S_i$	$a_i$	$b_i$	$S_i$
1	0.098	0.210	0.532	0.071	0.115	0.383
	4	2	0	3	7	6

Dalam penelitian ini juga dibuat aplikasi sederhana untuk menampilkan implementasi algoritma *K-Means*. Gambar 3, menunjukkan hasil cluster dari 10 data MKW yang telah dilakukan melalui sistem.



Gambar 3. Tampilan Aplikasi

#### 4. KESIMPULAN

Penerapan algoritma *K-Means* dalam menganalisa keterhubungan nilai yang diperoleh pada mata kuliah wajib (MKW) dan tugas akhir (TA) mahasiswa telah dipresentasikan. Hasil percobaan telah didapatkan 3 *cluster* berdasarkan MKW dengan nilai A, B, dan C, dimana C1 adalah sedikit, yang berhubungan dengan 1 MKW Pemrograman Visual; C2 adalah sedang, yang berhubungan dengan 6 MKW yaitu Basis Data, Sistem Keamanan Komputer, Pemrograman Berbasis Objek, Pemrograman Web, dan Struktur Data, sedangkan C3 adalah banyak, yang berkorelasi dengan 3 MKW yaitu Sistem Operasi, Sistem Keamanan Komputer, dan Kecerdasan Buatan. Perhitungan pengukuran hasil *cluster* menggunakan metode *Sum of Squared Errors* (SSE) dan evaluasi *cluster* menggunakan metode *Silhouette Coefficient* (SC), dimana, SSE 3 *cluster* bernilai 0.4506% dan 4 *cluster* bernilai 1.1072% telah didapatkan.

Hal ini menunjukkan bahwa 3 *cluster* lebih baik karena nilai kesalahan (*error*) yang diperoleh lebih kecil. Sedangkan, evaluasi *cluster* telah menggunakan SC menunjukkan bahwa 3 *cluster* memiliki rata-rata nilai 0.5852%. Hal ini berarti bahwa 3 *cluster* lebih baik dalam mengelompokkan MKW. Sedangkan, pada 4 *cluster* memiliki rata-rata nilai 0.4591% yang berarti bahwa cukup lemah dalam menempatkan objek dalam sebuah *cluster*. Hasil penelitian mengindikasikan bahwa metode *K-Means* dapat menjadi alternatif analisa hubungan nilai MKW dan TA sehingga diharapkan mahasiswa, dosen dan program studi dapat menjadikan bahan pendukung untuk membuat keputusan atau kebijakan dalam mengarahkan mahasiswa menentukan penelitian TA. Komparasi dan optimalisasi algoritma *K-Means* untuk menghasilkan berbagai akurasi menjadi penelitian selanjutnya.

#### 5. DAFTAR PUSTAKA

- Adrianto, R. (2016). Penerapan Metode Clustering Dengan Algoritma K-Means Untuk Rekomendasi Pemilihan Jalur Peminatan Sesuai Kemampuan Pada Progam Studi Teknik Informatika-S1 Universitas Dian Nuswantoro. *Journal JOINS Udinus*, 101–116.
- Anggara, M., Sujiani, H., & Helfi, N. (2016). Pemilihan Distance Measure Pada K-Means Clustering Untuk Pengelompokan Member Di Alvaro Fitness. *Jurnal Sistem Dan Teknologi Informasi*, 1(1), 1–6.
- Aroni, R. A. (2015). Penerapan Metode K-Means Untuk Clustering Mahasiswa Berdasarkan Nilai Akademik Dengan Weka Interface Studi Kasus Pada Jurusan Teknik Informatika UMM Magelang. *Ilmiah Semesta Teknik*, 18(1), 76–82.
- Farokhah, L., & Aditya, R. (2017). Implementasi K-Means Klustering Untuk Rekomendasi Tema Tugas Akhir Pada Stmik Asia Malang. *Jurnal Teknologi Dan Manajemen Informatika*, 3(2), 142–148.  
<https://doi.org/10.26905/jtmi.v3i2.1329>
- Gentle, J. E., Kaufman, L., & Rousseuw, P. J. (1991). Finding Groups in Data: An Introduction to Cluster Analysis. *Biometrics*.  
<https://doi.org/10.2307/2532178>
- Hasanah, N., Ugiarto, M., & Puspitasari, N. (2017). Sistem pengelompokan curah hujan menggunakan metode k-means di wilayah kalimantan timur. *Prosiding Seminar Nasional Ilmu Komputer Dan Teknologi Informasi*, 2(2), 122–126.
- Hermawan, T. N., Ugiarto, M., & Puspitasari, N. (2017). Sistem Evaluasi Kinerja Asisten Laboratorium Menggunakan Metode K-Means. *Prosiding Seminar Nasional Ilmu Komputer Dan Teknologi Informasi*, 3(2), 3–6.
- Kurnia Bakti, V., & Indriyatno, J. (2017). Klasterisasi Dokumen Tugas Akhir Menggunakan K-Means Clustering, Sebagai Analisa Penerapan Sistem Temu Kembali. *KOPERTIP : Jurnal Ilmiah Manajemen Informatika Dan Komputer*, 1(1), 31–34.  
<https://doi.org/10.32485/kopertip.v1i1.8>
- Kusuma, T. (2015). Pengelompokan Jenis Kupu-Kupu Menggunakan Fitur Ekstraksi GLCM dan Algoritma K-Means. *Jurnal Teknologi Informasi*, 1–4.
- Muttaqin, M. R., & Defriani, M. (2020). Algoritma K-Means untuk Pengelompokan Topik Skripsi Mahasiswa. *ILKOM Jurnal Ilmiah*.  
<https://doi.org/10.33096/ilkom.v12i2.542.121-129>
- Muzakir, A. (2014). Analisa Dan Pemanfaatan Algoritma K-Means Clustering Pada Data Nilai Siswa Sebagi Penentuan Penerima Beasiswa. *Prosiding Seminar Nasional Aplikasi Sains & Teknologi (SNAST)*, (November), 211–216.
- Parlina, I., Windarto, A. P., Wanto, A., & Lubis, M. R. (2018). Memanfaatkan Algoritma K-Means Dalam Menentukan Pegawai Yang Layak Mengikuti Asessment Center. *Memanfaatkan Algoritma K-Means Dalam Menentukan Pegawai Yang Layak Mengikuti Asessment Center Untuk Clustering Program Sdp*, 3(1), 87–93.
- Purnawansyah, & Haviluddin. (2017). K-Means clustering implementation in network traffic activities. *Proceedings - CYBERNETICSCOM 2016: International Conference on Computational Intelligence and Cybernetics*.  
<https://doi.org/10.1109/CyberneticsCom.2016.7892566>
- Purnawansyah, Haviluddin, Gafar, A. F. O., & Tahyudin, I. (2017). Comparison Between K-Means and Fuzzy C-Means Clustering in Network Traffic Activities. In *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*. [https://doi.org/10.1007/978-3-319-59280-0\\_24](https://doi.org/10.1007/978-3-319-59280-0_24)
- Puspitasari, N., & Haviluddin. (2016). Penerapan metode k-means dalam pengelompokan curah hujan di kalimantan timur. *PROCEEDINGS*

*SNRIK 2016.*

- Putra, G. M., Santosa, P., & Kusumawardani, S. (2015). Sistem Pendukung Keputusan Untuk Menentukan Prioritas Penerima Kartu Keluarga Sejahtera (KKS) Berdasar Tingkat Kemiskinan (Studi Kasus: Dinas Sosial Kota Samarinda). *Seminar Nasional Teknologi Terapan 2015 Sekolah Vokasi UGM*, 32–36. Yogyakarta: Sekolah Vokasi Universitas Gadjah Mada.
- Sari, I. P., & Devianto, D. (2014). Perancangan Sistem Pengenalan Wajah Manusia Menggunakan Web Camera dengan Metode Summary Squared Error ( SSE ). *Sistem Komputer*, 1–8.
- Siswanto, I., & Sampurna, Y. G. (2013). Faktor-Faktor Penghambat Dalam Pengerjaan Tugas Akhir Skripsi Mahasiswa FT UNY. *Journal Ilmiah*, 1–14.
- Sugiyama, M. (2015). Introduction to Statistical Machine Learning. In *Introduction to Statistical Machine Learning*. <https://doi.org/10.1016/C2014-0-01992-2>