

IMPLEMENTASI METODE *K-MEDOIDS CLUSTERING* UNTUK MENGETAHUI POLA PEMILIHAN PROGRAM STUDI MAHASISWA BARU TAHUN 2018 DI UNIVERSITAS KANJURUHAN MALANG

Bagus Wira¹, Alexius Endy Budianto², Anggri Sartika Wiguna³

Program Studi Teknik Informatika, Universitas Kanjuruhan^{1,2,3}

bgswira@gmail.com¹, endybudio@unikama.ac.id², anggrisartikawiguna@unikama.ac.id³

Abstrak. Setiap tahun Universitas Kanjuruhan Malang menerima hampir 2.000 mahasiswa yang tersebar diberbagai program studi. Oleh karenanya data yang telah ditampung pastinya banyak sekali, dari data tersebut dapat dilihat pola - pola pemilihan program studi berdasarkan nilai tes, asal sekolah, dan program studi. Penelitian ini menggunakan metode K-Medoids agar dapat diketahui pola pemilihan program studi bagi mahasiswa baru. K-Medoids merupakan metode *partisional clustering* dimana bertujuan untuk menemukan satu set *k-cluster* di antara data yang paling mencirikan objek dalam kumpulan suatu data. Hasil penelitian pengelompokan mahasiswa baru menunjukkan bahwa mahasiswa baru yang berasal dari SMA/SMK dengan nilai ujian diatas 70 mengambil jurusan TI, sedangkan mahasiswa baru yang berasal dari SMK dengan nilai ujian dibawah 70 dan SMA dengan nilai ujian dibawah 50 mengambil jurusan SI, dan sisanya mengambil jurusan NON TI/SI. Kualitas *cluster* yang dihasilkan berdasarkan proses pengujian yang dilakukan didapatkan hasil nilai *Silhouette Coefficient* terbaik yaitu 0.690754 dengan jumlah *cluster* sebanyak tiga dan jumlah data sebanyak 15.

Kata Kunci: *Clustering, K-Medoids, Silhouette Coefficient, Pola Pemilhan Program Studi.*

PENDAHULUAN

Banyaknya data yang tersimpan dalam *database* suatu universitas terkadang tidak dapat dimanfaatkan dengan baik selain sebagai kebutuhan administrasi, padahal setiap tahun terdapat penambahan data akibat penerimaan mahasiswa baru di lingkungan perguruan tinggi. Proses penerimaan mahasiswa baru menghasilkan data yang berlimpah melingkupi data pribadi mahasiswa tersebut, hal ini akan terus berlangsung setiap tahun sehingga data yang tersimpan akan semakin banyak dan sangat disayangkan jika data yang ada tidak dimanfaatkan dengan baik sebagai hal positif bagi universitas tersebut.

Sama halnya yang terjadi di lingkungan Universitas Kanjuruhan Malang, setiap tahun menerima hampir 2.000 mahasiswa yang terbagi dalam 17 program studi di enam fakultas untuk program S1. Sementara itu, untuk pasca sarjana ada dua program studi di dua fakultas. Oleh karenanya data yang telah ditampung pastinya banyak sekali, jika data tersebut dimanfaatkan dengan penerapan data mining maka akan dapat dilihat pola - pola pemilihan program studi berdasarkan nilai ujian, asal sekolah, dan prodi.

Mengingat pentingnya pengelompokan data mahasiswa baru, maka sangat perlu untuk dilakukan penelitian mengenai data mining, penelitian ini akan menggunakan metode *k-medoids* agar dapat diketahui pola pemilihan program studi bagi mahasiswa baru di lingkungan Universitas Kanjuruhan Malang. Sehingga dengan adanya penelitian ini dapat menjadi pengetahuan dan referensi baik bagi pengelola administrasi Universitas Kanjuruhan Malang maupun calon mahasiswa baru. Dari penelitian ini diharapkan dapat diketahui apakah mahasiswa tersebut benar - benar memilih program studi yang masih *linear* dengan bidang keilmuan yang dipelajari pada

jenjang pendidikan sebelumnya ataukah terjadi penyimpangan di SMA/SMK terdahulu terhadap pemilihan program studi yang ada di Universitas Kanjuruhan Malang.

2.1 Data Mining

Data mining adalah sebuah proses pencarian secara otomatis informasi yang berguna dalam tempat penyimpanan data berukuran besar. Istilah lain yang sering digunakan diantaranya *knowledge discovery (mining) in databae (KDD)*, *knowledge extraction*, data atau *pattern analysis*, data *archeology*, data *dredging*, *information harvesting*, dan *business intelligence*. Teknik data *mining* digunakan untuk memeriksa basis data berukuran besar sebagai cara untuk menemukan pola yang baru dan serbaguna. Tidak semua pekerjaan pencarian informasi dinyatakan sebagai data *mining*. Sebagai contoh, pencarian *record* individual menggunakan *database management system* atau pencarian halaman web tertentu melalui kueri ke semua *search engine* adalah pekerjaan pencarian informasi yang erat kaitannya dengan *information retrieval*. Teknik – teknik data *mining* dapat digunakan meningkatkan kemampuan sistem – sistem *information retrieval*.

2.2 Clustering

Menurut Larose (2015) *clustering* merupakan suatu proses pengelompokan record, observasi, atau mengelompokkan kelas yang memiliki kesamaan objek. Perbedaan *clustering* dengan klasifikasi yaitu tidak adanya variabel target dalam melakukan pengelompokan pada proses *clustering*. *Clustering* sering dilakukan sebagai langkah awal dalam proses data mining.

Kualitas hasil *clustering* sangat bergantung pada metode yang dipakai. Metode *clustering* juga harus dapat mengukur kemampuannya sendiri dalam usaha untuk menemukan suatu pola tersembunyi pada data yang sedang diteliti. Terdapat berbagai metode yang dapat digunakan untuk mengukur nilai kesamaan antar objek-objek yang dibandingkan, salah satunya ialah dengan *Euclidean Distance*. *Euclidean distance* menghitung jarak dua buah point dengan mengetahui nilai dari masing-masing atribut pada kedua poin tersebut. Berikut formula yang digunakan untuk menghitung jarak dengan *Euclidean Distance*:

$$Distance(p, q) = \left(\sum_k^n \mu_k |P_k - q_k|^r \right)^{1/r} \quad (2.1)$$

Keterangan:

- n = Jumlah record data
- k = Urutan field data
- r = pangkat 2
- μ_k = Bobot field yang diberikan *user*

2.3 K-Medoids Algorithm

2.3.1 Clustering K-Medoids

Algoritma *K-Medoids* atau sering disebut juga dengan algoritma PAM (*Partitioning Around Medoid*) dikembangkan oleh Leonard Kaufman dan Peter J. Rousseeuw, dimana merupakan algoritma yang mirip dengan *k-means* karena kedua algoritma ini *partitional* yang memecah dataset menjadi kelompok – kelompok. Perbedaan dari algoritma *k-means* dengan algoritma *K-Medoids* terletak pada penentuan pusat *cluster*, dimana algoritma *k-means* menggunakan nilai rata – rata (*means*) dari setiap *cluster* sebagai pusat *cluster* dan algoritma *K-Medoids* menggunakan objek data sebagai perwakilan (*medoids*) sebagai pusat *cluster* (Kaur, dkk., 2014). Algoritma *K-Medoids* digunakan untuk mengatasi kelemahan dari algoritma *k-means* yang sangat sensitif terhadap pencilaan (*outlier*) karena objek – objek ini sangat jauh letaknya/karakteristiknya dari

mayoritas data lainnya, sehingga jika dimasukkan ke suatu *cluster* data semacam ini bisa mendistorsi nilai rata-rata (*mean*) *cluster* tersebut.

2.3.2 Langkah Clustering K-Medoids

Algoritma *K-Medoids* merupakan teknik partisi klasik dari *clustering* yang melakukan klasterisasi dataset objek n ke dalam k *cluster* yang dikenal sebagai *a priori* (Abhishek & Purnima, 2013). Algoritma ini beroperasi pada prinsip untuk meminimalkan jumlah kesamaan antara setiap objek dan titik referensi yang sesuai. Algoritma *K-Medoids* dapat dilakukan dengan langkah – langkah sebagai berikut (Bhat, 2014):

- a. Inisialisasi pusat *cluster* sebanyak k (jumlah *cluster*).
- b. Hitung setiap objek ke *cluster* terdekat menggunakan persamaan ukuran jarak *Euclidian Distance*. Perhitungan *Euclidian Distance* menggunakan persamaan (2.1)
- c. Setelah menghitung jarak *Euclidian Distance*, inisialisasikan pusat *cluster* baru secara acak pada masing – masing objek sebagai kandidat *non medoids*.
- d. Hitung jarak setiap objek yang berada pada masing – masing *cluster* dengan kandidat *non medoids*.
- e. Hitung total simpangan (S) dengan menghitung total *distance* baru – total *distance* lama. Jika $S < 0$ maka tukar objek dengan data *cluster non medoids* untuk membentuk sekumpulan k objek baru sebagai *medoids*.
- f. Ulangi langkah c – e hingga tidak terjadi perubahan pada *medoid*, sehingga di dapatkan *cluster* beserta anggota *cluster* masing – masing.

2.4 Silhouette Coefficient

Silhouette Coefficient merupakan sebuah metode yang digunakan untuk melihat kualitas dan kekuatan *cluster*, seberapa baik suatu objek ditempatkan dalam suatu *cluster*. Metode *Silhouette Coefficient* ini merupakan sebuah metode gabungan dari metode *cohesion* dan metode *separation*. Metode *cohesion* adalah sebuah ukuran seberapa dekat relasi antara objek dalam sebuah *cluster*. Sedangkan metode *separation* adalah sebuah ukuran seberapa jauh atau terpisah sebuah *cluster* dengan *cluster* yang lainnya.

Tahapan perhitungan *Silhouette Coefficient* antara lain [Handoyo, dkk., 2014]:

1. Hitung jarak rata-rata dari suatu objek misalkan objek ke- i dengan semua objek lain yang berada di dalam satu *cluster*.

$$a(i) = \frac{1}{[A] - 1} \sum_{j \in A, j \neq i} d(i, j) \quad (2.2)$$

2. Hitung rata-rata jarak dari objek ke- i tersebut dengan semua objek pada *cluster* lainnya, kemudian ambillah nilai terkecilnya.

$$d(i, C) = \frac{1}{[A]} \sum_{j \in C} d(i, j) \quad (2.3)$$

3. Nilai *Silhouette Coefficient* nya adalah:

Jumlah $s(i)$ diperoleh dengan menggabungkan $a(i)$ dan $b(i)$:

$$s \begin{cases} 1 - a(i)/b(i) & \text{if } a(i) < b(i), \\ 0 & \text{if } a(i) = b(i) \\ b(i)/a(i) - 1 & \text{if } a(i) > b(i) \end{cases}$$

Sehingga dapat dirumuskan:

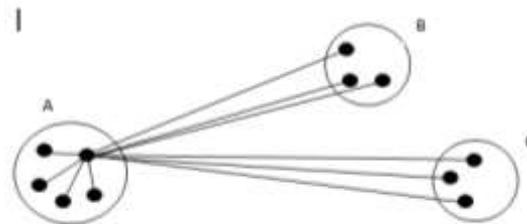
$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (2.4)$$

Nilai hasil perhitungan dengan menggunakan metode *Silhouette Coefficient* terletak pada kisaran antara nilai -1 hingga 1. Nilai rata-rata *Silhouette Coefficient* dari setiap objek dalam suatu *cluster* merupakan suatu ukuran yang menunjukkan seberapa dekat kemiripan data dikelompokkan didalam satu *cluster* tersebut. Semakin nilai rata-rata *Silhouette Coefficient* mendekati nilai 1, maka semakin baik pengelompokan data di dalam satu *cluster*. Sebaliknya jika nilai rata-rata *Silhouette Coefficient* mendekati -1, maka semakin buruk pengelompokan data di dalam satu *cluster*. Berikut kriteria pengukuran nilai *Silhouette Coefficient* [Rousseeuw, 1987].

Tabel 1. Ukuran Nilai *Silhouette Coefficient*

<i>Silhouette Coefficient</i>	Interpretasi yang diusulkan
$0.7 < SC \leq 1.0$	<i>Stronge Structure</i>
$0.5 < SC \leq 0.7$	<i>Mediun Structure</i>
$0.25 < SC \leq 0.5$	<i>Weak Structure</i>
$SC \leq 0.25$	<i>No Structure</i>

Berikut merupakan gambar ilustrasi dari *silhouette coefficient*:



Gambar 1. Ilustrasi elemen-elemen yang termasuk dalam perhitungan $s(i)$, dimana objek data ke- i termasuk dalam cluster A (Rousseeuw, 1987)

METODE PENELITIAN

Metode penelitian menjelaskan langkah – langkah yang akan digunakan serta perancangan dalam melakukan implementasi. Metode *K-Medoids clustering* untuk pengelompokkan data pola pemilihan program studi mahasiswa baru.

Gambar 3.1 menunjukkan alur dari diagram blok metodologi penelitian dimana pada setiap tahapnya memiliki fungsi serta tugas masing – masing yang saling berkaitan pada satu tahapan dengan tahapan berikutnya.



Gambar 2. Diagram Blok Metode Penelitian

3.1 Studi Literatur

Dalam penelitian ini, studi literatur berfungsi untuk menjelaskan dasar teori dari berbagai sumber acuan yang digunakan dalam penulisan penelitian *clustering* untuk pengelompokan data mahasiswa baru menggunakan metode *K-Medoids Clustering*

3.2 Pengumpulan Data

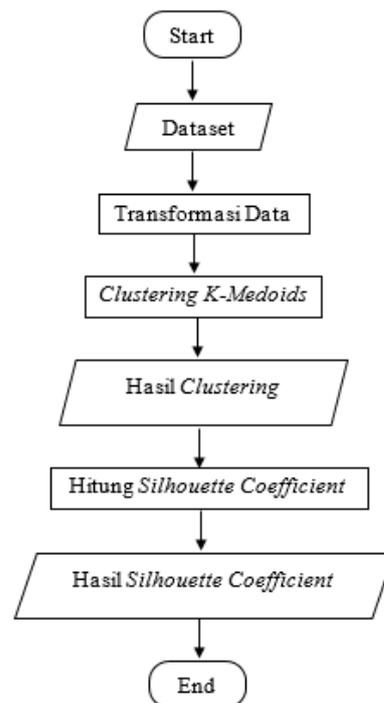
Pada tahap ini dilakukan pengumpulan data dari objek yang akan diteliti dimana data ini diperoleh dari Universitas Kanjuruhan Malang. Data yang digunakan memiliki 3 variabel yang berisikan nilai ujian tes masuk kuliah, asal sekolah, dan program studi yang dipilih.

3.3 Analisis Kebutuhan

Analisis kebutuhan dilakukan untuk menentukan kebutuhan apa saja yang dibutuhkan dalam penelitian yang akan dilakukan, kebutuhan ini akan digunakan dalam proses membangun perangkat lunak.

3.4 Perancangan Sistem

Pada tahap perancangan sistem ini terdapat beberapa deskripsi sistem yang akan dibuat secara umum yaitu perancangan aplikasi, manualisasi, rancang antar muka aplikasi, pengujian dan analisis.



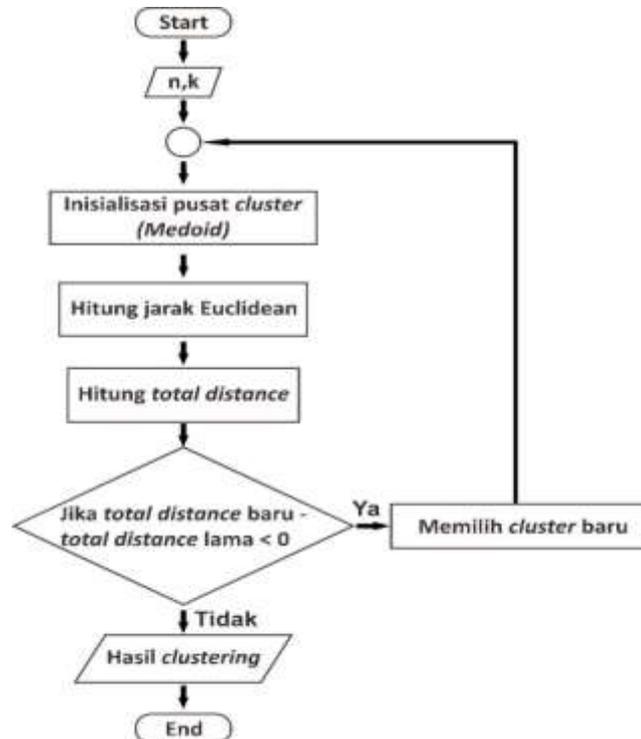
Gambar 2. Flowchart Perancangan Perangkat Lunak

Berikut merupakan penjelasan tahapan - tahapan dalam proses perancangan perangkat lunak pada penerapan metode *K-Medoids Clustering* untuk pengelompokan data mahasiswa baru:

- a. Melakukan input data berupa data mahasiswa baru.
- b. Melakukan transformasi data yang telah di inputkan untuk proses perhitungan selanjutnya.
- c. Melakukan proses klasterisasi dengan menggunakan algoritma *K-Medoids*.
- d. Melakukan proses validasi terhadap kualitas *cluster* dengan menghitung nilai *Silhouette Coefficient*

- e. Menampilkan keluaran hasil dari *clustering* dengan kualitas *cluster* yang dapat dilihat dari nilai *Silhouette Coefficient* yang dihasilkan.

Secara garis besar proses metode *K-Medoids* ditunjukkan pada diagram alir pada Gambar.3.



Gambar 3. Diagram Alir Proses *K-Medoids Clustering*

3.5 Implementasi dan Pengujian Sistem

Dalam hal ini implementasi merupakan tahap dimana sistem siap dioperasikan pada keadaan yang sebenarnya, sehingga akan diketahui apakah sistem yang dibuat benar-benar dapat menghasilkan tujuan yang ingin dicapai. Sistem ini dibangun dengan berbasis *web* dengan menggunakan bahasa pemrograman PHP. Pada pengujian perangkat lunak akan dilakukan pengujian terhadap jumlah *cluster* dan jumlah data terbaik. Teknik pengujian ini dilakukan menggunakan metode *black box* dengan mengevaluasi cara kerja sistem yang telah dibuat.

HASIL DAN PEMBAHASAN

4.1 Analisis Data

Secara umum, sistem yang akan dibangun dalam penelitian ini adalah sebuah sistem dengan fungsi utama untuk melakukan pengelompokkan calon mahasiswa baru. Data yang digunakan untuk penelitian merupakan data yang diperoleh dari Penerimaan Mahasiswa Baru (PMB) tahun 2018 di Universitas Kanjuruhan Malang. Adapun pengelompokkan data tersebut digunakan langkah – langkah sebagai berikut:

1. Sumber Data

Sumber data dalam penelitian ini langsung diambil dari Universitas Kanjuruhan Malang dimana data tersebut merupakan data sekunder yang terdiri atas data mahasiswa baru tahun ajaran 2018 – 2019 sebanyak 15 data yang sudah di sorting.

2. Transformasi Data

Pada tahap ini dilakukan proses transformasi data tujuannya adalah untuk menyamakan skala atribut data dalam range yang lebih spesifik, sehingga data dapat diolah dengan mudah menggunakan metode *K-Medoids Clustering*. Variabel yang dipilih pada data pendaftaran calon mahasiswa baru yaitu nilai ujian, asal sekolah, dan program studi yang dipilih. Untuk variabel nilai ujian dikelompokkan menjadi 3 kelompok, yang pertama untuk nilai ujian dengan rata – rata ≤ 50 di transformasikan dengan nilai 1, >50 dan data ≤ 70 di transformasikan dengan nilai 2 dan > 70 di transformasikan dengan nilai 3. Untuk variabel asal sekolah di kelompokkan menjadi 2 kelompok yang pertama untuk asal sekolah SMA di transformasikan dengan nilai 1 dan asal sekolah SMK di transformasikan dengan nilai 2.

3. Pengolahan Data

Setelah proses transformasi langkah selanjutnya adalah proses pengolahan data menggunakan metode *K-Medoids Clustering*. Tahapan proses metode *K-Medoids Clustering* adalah sebagai berikut :

- a. Ditentukan k (jumlah *cluster*) dari n objek adalah 3
- b. Tentukan *cluster* awal sebagai *medoid* dengan asumsi seperti pada Tabel 4.1 berikut

Tabel 2. Inisialisasi *medoids*

Nama	Keterangan	x	y	z
C1	Diambil data ke-6 sebagai pusat cluster ke-1	1	2	3
C2	Diambil data ke-8 sebagai pusat cluster ke-2	2	1	1
C3	Diambil data ke-14 sebagai pusat cluster ke-3	1	2	1

- c. Tempatkan objek – objek *non medoids* ke dalam *cluster* yang paling dekat dengan *medoids* berdasarkan jarak *Euclidean*. Berikut contoh perhitungan jarak pada data ke 1:

$$d(x_1c_1) = \sqrt{(3 - 1)^2 + (2 - 2)^2 + (1 - 3)^2} = 2,828427$$

$$d(x_1c_2) = \sqrt{(3 - 2)^2 + (2 - 1)^2 + (1 - 1)^2} = 1,414214$$

$$d(x_1c_3) = \sqrt{(3 - 1)^2 + (2 - 2)^2 + (1 - 1)^2} = 2$$

Cara yang sama dapat dilakukan untuk perhitungan jarak data ke-2 sampai data ke-15, sehingga nanti hasil keseluruhannya datanya akan diperoleh seperti pada tabel 4.2 berikut.

Tabel 3. Hasil perhitungan jarak ke setiap medoids

Data ke i	C ₁	C ₂	C ₃	Jarak
1	2,828	1,414	2	2
2	2	2,449	2,828	1
3	2	1,414	0	3
4	2,449	1,414	2,449	2
5	2,449	1,414	2,449	2
6	0	2,449	2	1
7	1	2,236	2,236	1
8	2,449	0	1,414	2
9	2,236	1,732	2,236	2
10	3	1	2,236	2
11	1	2,236	2,236	1
12	2,449	0	1,414	2
13	2,236	1	1	3
14	2	1,414	0	3
15	1	1,732	1	3
JUMLAH	29,09852	21,90628	25,50011	
Total Cost		76,50491		

d. Setelah itu, tentukan objek *non medoid* dengan asumsi sebagai berikut.

Tabel 4. Inisialisasi non medoids

Nama	Keterangan	x	y	z
D1	Diambil data ke-5 sebagai pusat cluster ke-1	3	1	2
D2	Diambil data ke-11 sebagai pusat cluster ke-2	1	1	3
D3	Diambil data ke-13 sebagai pusat cluster ke-3	2	2	1

e. Ulangi langkah c untuk objek *non medoids*

$$d(x_1c_1) = \sqrt{(3 - 3)^2 + (2 - 1)^2 + (1 - 1)^2} = 1,414214$$

$$d(x_1c_2) = \sqrt{(3 - 1)^2 + (2 - 1)^2 + (1 - 2)^2} = 3$$

$$d(x_1c_3) = \sqrt{(3 - 2)^2 + (2 - 2)^2 + (1 - 1)^2} = 1$$

f. Sehingga hasilnya diperoleh seperti pada tabel 4.4 berikut.

Tabel 5. Hasil perhitungan jarak ke setiap non medoids

Data ke /	C ₁	C ₂	C ₃	Jarak
1	1,414	3	1	3
2	1,414	2,236	2,236	1
3	2,449	2,236	1	3
4	0	2,236	1,732	1
5	0	2,236	1,732	1
6	2,449	1	2,236	2
7	1,732	1,414	2	2
8	1,414	2,236	1	3
9	1	2,449	1,414	1
10	1	2,828	1,414	1
11	2,236	0	2,449	2
12	1,414	2,236	1	3
13	1,732	2,449	0	3
14	2,244	2,236	1	3
15	2,236	1,414	1,414	2
JUMLAH	22,94156	30,20831	21,62837	
Total				74,77824
Distance				

- g. Hitung total simpangan (*S*) dengan menghitung total *distance* baru – total *distance* lama. Jika $S < 0$ maka tukar objek dengan data *cluster* untuk membentuk sekumpulan *k* objek baru sebagai *medoid* (Setyawati, 2017).

$$S = \text{total distance baru} - \text{total distance lama}$$

$S = 76,79409 - 74,77824 = -$

- h. Karena nilai $S < 0$ maka proses pengklasteran dihentikan. Sehingga objek *non medoids* dijadikan sebagai *medoid* baru.
- i. Sehingga *cluster* yang di dapatkan seperti pada tabel 4.5.

Tabel 6. Hasil pengclusteran mahasiswa baru dengan k-medoids

Data ke i	Nilai Ujian (x)	Asal Sekolah (y)	Prodi (z)	Cluster yang
				dikuti
1	75	SMK	TI	3
2	88	SMK	NON	1
3	44,5	SMK	TI	3
4	72	SMA	SI	1
5	78	SMA	SI	1
6	35,4	SMK	NON	2
7	66,6	SMK	NON	2
8	56,5	SMA	TI	3
9	82,1	SMK	SI	1
10	87	SMA	TI	1
11	27,3	SMA	NON	2
12	55	SMA	TI	3
13	66	SMK	TI	3
14	32	SMK	TI	3
15	49,2	SMK	SI	2

- j. Setelah mendapatkan hasil *cluster* akhir, hitung nilai *average dissimilarity* (a_i) untuk evaluasi dengan metode *silhouette coefficient*. Perhitungan $a(i)$ menggunakan persamaan (2.2)

$$a(i) = \sqrt{(1 - 1,414214)^2} = \sqrt{0,171573} = 0,414214$$

$$a(i) = \sqrt{(1 - 0)^2} = \sqrt{1} = 1$$

$$a(i) = \sqrt{(1 - 0)^2} = \sqrt{1} = 1$$

$$a(i) = \sqrt{(1 - 1)^2} = \sqrt{0} = 0$$

$$a(i) = \sqrt{(1 - 1)^2} = \sqrt{0} = 0$$

$$= (0,41424 + 1 + 1 + 0 + 0) / 5 = 0,482843$$

- k. Lalu hitung nilai *lowest average dissimilarity* (b_i). Perhitungan $b(i)$ menggunakan persamaan (2.3)

$$(0,482843) \rightarrow (1) = \text{distance}$$

$$= \sqrt{(1 - 1)^2 + (1,414214 - 1)^2 + (0 - 1)^2 + (0 - 1)^2 + (1 - 1)^2 + (1 - 1)^2} = 1,473626$$

$$= (1,473626 + 2,084022 + 2 + 2,084022) / 4 = 1,910417$$

$$(0,482843) \rightarrow (1) = \text{distance}$$

$$= \sqrt{(1 - 1)^2 + (1,414214 - 1)^2 + (0 - 1)^2 + (0 - 1)^2 + (1 - 1)^2 + (1 - 1)^2} = 1,473626$$

$$= (1,473626 + 1,473626 + 1,473626 + 1,473626 + 2 + 1,473626) / 6 = 1,561355$$

Nilai minimum antara kedua *cluster* ialah $b(i) = (1,561355 < 1,910417) = 1,561355$

- l. Menghitung nilai *silhouette coefficient* (s_i). Perhitungan $s(i)$ menggunakan persamaan (2.4)

$$s(i) = 1 - \left(\frac{a_i}{b_i}\right) = 1 - \left(\frac{0,482843}{1,561355}\right) = 0,690754$$

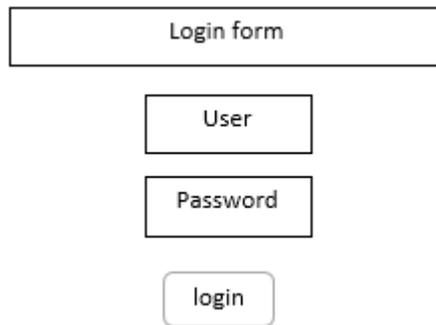
- m. Karena hasil nilai *silhouette coefficient* mendekati nilai 1 maka pengelompokan data dalam *cluster* dinyatakan baik.

Dari perhitungan dan pengelompokan data diatas dapat di ambil kesimpulan, bahwa rata – rata mahasiswa baru yang berasal dari sekolah SMA dan SMK dengan nilai > 70 masuk ke dalam *cluster* 1, sedangkan untuk mahasiswa baru yang berasal dari SMA dan SMK dengan nilai < 70 masuk ke dalam *cluster* 2 dan 3.

Jadi, itu artinya mahasiswa baru yang berasal dari sekolah SMA dan SMK yang nilai ujiannya diatas 70 mengambil jurusan TI, sebaliknya mahasiswa baru yang berasal dari sekolah SMA dan SMK yang nilai ujiannya dibawah 70 mengambil jurusan SI atau diluar SI dan TI.

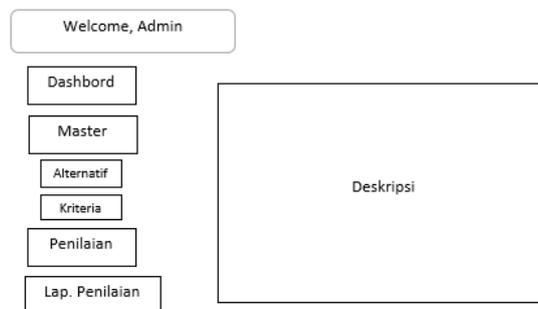
4.2 Rancangan *Interface Program*

Di dalam sebuah membangun sistem tentunya sebuah rancangan sangat diperlukan agar saat proses pembuatannya dapat di mengerti dengan mudah.



Gambar 4. Rancangan *interface login*

Digambar perancangan diatas setelah admin melakukan login, akan langsung masuk ke tampilan utama.



Gambar 5. Tampilan utama

Pada perancangan tampilan utama, admin akan menginput data pada menu master dan memilih sub menu alternatif. Untuk sub menu kriteria nantinya akan berisikan tentang variabel pada pemilihan program studi, mulai dari nilai ujian, asal sekolah, dan prodi yang dipilih.



Gambar 6. Rancangan input data

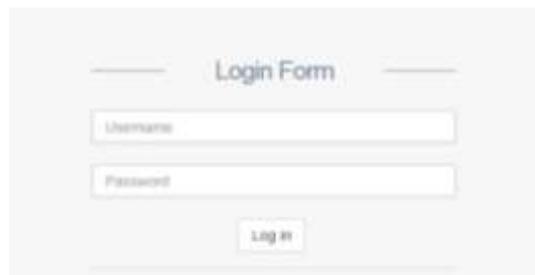
Diperancangan ini nantinya berfungsi untuk menginputkan data mahasiswa baru berdasarkan hasil penelitian di Universitas Kanjuruhan Malang.

The image shows a web form for grading. It consists of four input fields stacked vertically: 'Nama', 'Nilai Ujian' (with a '--' placeholder), 'Asal Sekolah' (with a '--' placeholder), and 'Prodi' (with a '--' placeholder). Below these fields are two buttons: 'Save' and 'Reset'.

Gambar 7. Rancangan penilaian

Pada perancangan menu penilaian ini nantinya akan di inputkan variabel yang akan di hitung berdasarkan data mahasiswa baru yakni nilai ujian, asal sekolah, dan prodi, kemudian dilakukan perhitungan secara otomatis menggunakan metode *k-medoids clustering*.

4.3 Implementasi dan Pengujian



Gambar 8. Tampilan awal

Tabel 7. Deskripsi pengujian tampilan awal

Kebutuhan	Skenario	Hasil yang di harapkan	Hasil pengamatan	Hasil pengujian
Pengujian tampilan awal	Saat admin tidak menginputkan password	Harap isi password terlebih dahulu	Sesuai	Sesuai
	Saat admin tidak menginputkan username	Harap isi username terlebih dahulu	Sesuai	Sesuai

Gambar 9. Tampilan sub menu alternatif untuk inputan data mahasiswa baru

Tabel 8. Deskripsi pengujian sub menu alternatif inputan data

Kebutuhan	Skenario	Hasil yang di harapkan	Hasil pengamatan	Hasil pengujian
Pengujian interface pada sub menu alternatif untuk penginputan data mahasiswa baru	Saat admin tidak menginputkan nama mahasiswa baru	Harap terlebi h dahulu	isi Sesuai	Sesuai
	Saat admin tidak menginputkan jenis kelamin mahasiswa baru	Harap terlebi h dahulu	isi Sesuai	Sesuai
	Saat admin tidak menginputkan tempat lahir mahasiswa baru	Harap terlebi h dahulu	isi Sesuai	Sesuai
	Saat admin tidak menginputkan tanggal lahir mahasiswa baru	Harap terlebi h dahulu	isi Sesuai	Sesuai

Gambar 10. Tampilan menu penilaian

Gambar 11. Tampilan inputan variabel mahasiswa baru pada menu penilaian

Tabel 9. Deskripsi pengujian penilaian

Kebutuhan	Skenario	Hasil yang di harapkan	Hasil pengamatan	Hasil pengujian
Pengujian interface pada menu penilaian untuk inputan nilai variabel mahasiswa baru	Saat admin tidak menginputkan variabel nilai ujian	Harap terlebih dahulu	isi Sesuai	Sesuai
	Saat admin tidak menginputkan variabel asal sekolah	Harap terlebih dahulu	isi Sesuai	Sesuai
	Saat admin tidak menginputkan variabel prodi	Harap terlebih dahulu	isi Sesuai	Sesuai

Setelah di inputkan semua variabel pada penilaian, maka akan langsung menghitung secara otomatis nilai variabel yang di inputkan menggunakan metode *K-Medoids Clustering*, hasilnya akan tampak seperti gambar 4.18

No	ID alternatif	K1 (nilai ujian)	K2 (asal sekolah)	K3 (prodi)	Cluster 1 (T)	Cluster 2 (S)	Cluster 3 (Non T/S)	Hasil
11	ALT1011	3	2	3	2	2.445	2.878	Cluster 1 (T)
Jumlah:					29.192	18.694	26.540	Total Cost 59.400

Gambar 12. Tampilan menu laporan penilaian

PENUTUP

Berdasarkan hasil penelitian yang telah dilakukan dan dibahas pada bab tiga dan empat, dapat diambil kesimpulan sebagai berikut:

Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan dapat diambil kesimpulan sebagai berikut:

1. Algoritma *K-Medoids* dapat di implementasikan untuk proses pola pemilihan program studi mahasiswa baru sesuai dengan tahapan – tahapan perhitungan *K-Medoids Clustering*.
2. Pengelompokan mahasiswa baru menggunakan metode *K-Medoids Clustering* menunjukkan bahwa mahasiswa baru yang berasal dari sekolah SMA/SMK dengan nilai ujian diatas 70 mengambil jurusan TI, sedangkan mahasiswa baru yang berasal dari sekolah SMK dengan nilai ujian dibawah 70 dan SMA dengan nilai ujian dibawah 50 mengambil jurusan SI, dan sisanya mengambil jurusan NON TI/SI.
3. Kualitas *cluster* yang dihasilkan berdasarkan proses pengujian yang dilakukan didapatkan hasil nilai *Silhouette Coefficient* terbaik yaitu 0.690754 dengan jumlah *cluster* 3 dan jumlah data 15. Berdasarkan hasil tersebut dapat diambil kesimpulan bahwa jumlah *cluster* dan jumlah data mempengaruhi terhadap hasil kualitas dari *cluster*.

Saran

Saran penulisan terkait penelitian “Implementasi Metode *K-Medoids Clustering* Untuk Mengetahui Pola Pemilihan Program Studi Mahasiswa Baru di Universitas Kanjuruhan Malang” untuk penelitian selanjutnya sebagai berikut:

1. Dianjurkan melakukan penambahan jumlah data dan parameter, dimana dalam penentuan parameter dilakukan peninjauan ulang dengan menggunakan bantuan sebuah algoritma atau dengan bantuan para ahli untuk prosedur pemilihan parameter, sehingga tingkat kualitas *cluster* yang dihasilkan lebih baik.
2. Penelitian selanjutnya dapat mengembangkan perangkat lunak dengan menggunakan metode *clustering* lain, seperti metode *DBSCAN* untuk dapat membandingkan nilai kualitas *cluster* yang terbentuk dengan hasil dari penelitian sebelumnya.

DAFTAR PUSTAKA

- Andre. 2014. Tutorial Belajar CSS: Cara Mengatur Tinggi Baris dalam Paragraf HTML (line-height), (Online), (<https://www.duniailkom.com/tutorial-belajar-css-cara-mengatur-tinggi-baris-dalam-paragraf-html-line-height/>), diakses 22 Januari 2019.
- Chrisnanto, Yulison H. & Abdillah, G. 2015. Penerapan Algoritma Partitioning Around Medoids (PAM) Clustering Untuk Melihat Gambaran Umum Kemampuan Akademik Mahasiswa. Seminar Nasional Teknologi Informasi dan Komunikasi, (Online), 15 (1): 444-448, (<https://fti.uajy.ac.id/sentika/publikasi/makalah/2015/56.pdf>), diakses 21 Maret 2018.
- Defiyanti, Sofi., Jajuli, Mohamad, & Nurul R. W. 2017. Optimalisasi K-Medoid Dalam Pengklasteran Pelamar Beasiswa Dengan Cubic Clustering Criterion, Teknosi, (Online), 17(01): 211-218, (<http://teknosi.fti.unand.ac.id/index.php/teknosi/article/view/222>), diakses 30 Januari 2019.
- Han, J., Pei, J., and Kamber, M. 2011. Data Mining: Concepts and Techniques. Third Edition, Morgan Kaufman.
- Irwansyah, E., and Faisal M. 2015. Advanced Clustering Teori dan Aplikasi. Yogyakarta: DeePublish.
- Irwansyah, Edy, S.T., M.SI. 2017. Clustering, (Online), (<http://socs.binus.ac.id/2017/03/09/clustering/>), diakses pada 7 Februari 2019.
- Kadir, Abdul. 2017. Dasar Logika Pemrograman Komputer. Jakarta: Elex Media Komputindo
- Kaur, Noor K., Kaur, Usvir. & Singh, Dr. Dheerendra. 2014. K-Medoid Clustering Algorithm- A Review, International Journal of Computer Application and Technology, (Online), 14 (1):

42-45, (https://www.academia.edu/8446443/K-Medoid_Clustering_Algorithm-A_Review?auto=download), diakses 22 Maret 2018.

Kusrini, and Luthfi E.T. 2009. Algoritma Data Mining. Yogyakarta: Andi.

Larose, Daniel, T. and Larose, Chantal D. 2015. Data Mining and Predictive Analytics. Second Edition, John Wiley & Sons.

Pramesti, Dyang F., Furqon M. Tanzil, & Dewi, C. 2017. Implementasi Metode K-Medoids Clustering Untuk Pengelompokan Data Potensi Kebakaran Hutan/Lahan Berdasarkan Persebaran Titik Panas (Hotspot), Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, (Online), 17 (9): 723-732, (<http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/download/204/101/>), diakses 23 Februari 2018.

Solichin, Ahmad. 2016. Pemrograman Web dengan PHP dan MySQL. Jakarta: Universitas Budi Luhur.