

Gambaran Umum Kemampuan Akademik Mahasiswa Unjani Dengan Algoritma *Partitioning Around Medoids* (PAM) Clustering

Yulison Herry Chrisnanto, Gunawan Abdillah

Jurusan Informatika, Fakultas MIPA

Universitas Jenderal Achmad Yani

Jl. Terusan Sudirman, Cimahi

y.chrisnanto@gmail.com

Abstrak—Kajian terkait kedataan sudah banyak dilakukan, salah satu kajian kedataan yang menarik adalah eksplorasi data transaksi dalam jumlah yang besar, diproses sedemikian rupa untuk diekstrak menjadi pulau-pulau informasi yang sebelumnya belum diketahui serta memiliki nilai yang tinggi bagi kebutuhan analisis dalam rangka pengambilan keputusan. Kajian kedataan ini dikenal sebagai *data mining*. Penelitian ini akan melakukan kajian kedataan, khususnya data akademik mahasiswa di Unjani angkatan 2013/2014, semester ganjil dan genap. Pendekatan metodologi yang digunakan adalah *Clustering Data Mining* untuk mendapatkan gambaran umum kemampuan akademik mahasiswa Unjani. Teknik Clustering yang akan digunakan adalah *Partitioing Around Medoids* (PAM), yaitu pendekatan dalam mempartisi data sesuai dengan kedekatan karakteristik objek data dalam dataset akademik. Hasil akhir dari penelitian ini berupa produk perangkat lunak komputer yang akan memproses data akademik mahasiswa Unjani menjadi informasi yang memberikan gambaran umum kemampuan akademik mahasiswa di Unjani.

Kata kunci—*clustering, PAM, Medoids, Data Mining*

I. PENDAHULUAN

Unjani merupakan institusi pendidikan tinggi dimana untuk menghadapi persaingan saat ini perlu meningkatkan kemampuan pengelolaan proses akademik secara berkesinambungan, agar lulusannya dapat diterima oleh para pemangku kepentingan yang lebih luas, khususnya menjelang era perdagangan bebas, dimana kebutuhan tenaga profesional secara regional sudah mulai diberlakukan. Oleh karena itu kemampuan akademik sesuai dengan bidang yang digeluti harus semakin dapat bersaing, baik nasional maupun regional.

Untuk dapat mengetahui kemampuan akademik seorang mahasiswa secara umum, diperlukan suatu mekanisme kalkulasi yang melibatkan banyak aspek akademik, dan hal tersebut menjadi bagian paling sulit ketika data akademik tersebar dalam beragam format data. Oleh karena itu proses terhadap kedataan yang terkait dengan data akademik mahasiswa menjadi kajian menarik untuk dilakukan penelitiannya, dimana model ekstraksi data akademik mahasiswa menjadi pengetahuan dapat dilakukan untuk melalui teknik-teknik tertentu, salah satunya adalah dengan teknik clustering. Penelitian ini diarahkan pada ekstraksi data akademik mahasiswa menggunakan algoritma clustering.

II. METODE

A. *Data Mining*

Secara sederhana, data mining merujuk pada ekstraksi atau "penambangan" pengetahuan dari sejumlah besar data. Istilah "panambangan" sebenarnya kurang tepat, data mining seharusnya lebih tepat bernama "penambangan pengetahuan dari data," Langkah-langkah penting dalam proses penambangan pengetahuan dari data secara umum sebagai berikut : Data cleaning, yaitu proses untuk menghilangkan data yang dipandang tidak konsisten atau tidak diperlukan, Data integration, dimana data yang berasal dari berbagai sumber akan digabungkan, Data selection, adalah data yang dipilih sesuai dengan kebutuhan analisis, Data transformation, merupakan proses transformasi data ke dalam bentuk yang siap untuk ditambang, Data mining, merupakan proses esensial dimana metoda kecerdasan diaplikasikan untuk mengekstraksi data. Pattern evaluation, untuk mengidentifikasi pola-pola yang menarik yang merepresentasikan pengetahuan dan terakhir

adalah *knowledge presentation*, merupakan teknik untuk memvisualisasikan pengetahuan kepada pemakai [1]. Proses ekstraksi pengetahuan dalam data mining menggunakan teknik-teknik tertentu, antara lain : statistik, machine learning dan pattern recognition. Teknik clustering sering digunakan dalam proses data mining, dimana teknik ini dapat dikelompokkan dalam empat katagori, antara lain : *partition based*, *hierarchical based*, *density based* dan *grid based* [7].

Definisi dari data mining dapat dijelaskan sebagai berikut [6] :

- Data mining atau *knowledge discovery in databases*, juga dikenal sebagai *nontrivial extraction*, secara implisit merupakan proses pemanfaatan data yang memiliki potensi kebergunaan tinggi yang sebelumnya tidak diketahui dengan pendekatan sejumlah teknik seperti clustering, *data summarization*, klasifikasi, *finding dependency network* serta deteksi anomali-anomali.
- Data mining mengacu pada menggunakan berbagai teknik untuk mengidentifikasi sebarang informasi atau pengetahuan dalam basis data dalam rangka pengambilan keputusan yang diterapkan di berbagai bidang seperti pendukung keputusan, prediksi, peramalan dan estimasi. Data sering berjumlah besar, akan tetapi memiliki nilai yang rendah sehingga perlu dilakukan penambangan untuk menemukan informasi yang tersembunyi dalam basis data.
- Data mining adalah proses menemukan informasi bermakna melalui pola korelasi baru dan trend dengan memilah-milah sejumlah besar data yang disimpan dalam repositori dengan menggunakan teknik pengenalan pola serta teknik statistik dan matematika.

Dalam prakteknya, data mining kebanyakan memiliki tujuan dalam melakukan prediksi dan deskripsi, yaitu :

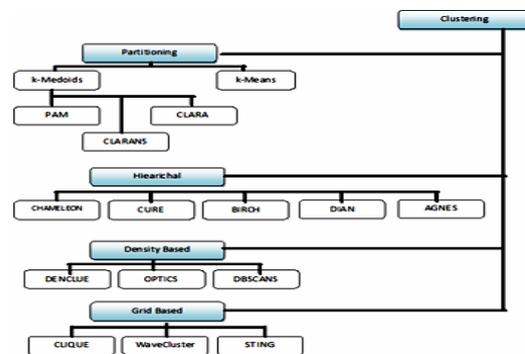
- Prediksi, yaitu memanfaatkan variabel yang ada dalam basis data untuk memprediksi nilai yang tidak diketahui atau sesuatu yang menarik untuk masa yang akan datang.
- Deskripsi berfokus pada menemukan pola yang menggambarkan data dan selanjutnya presentasi untuk interpretasi pengguna.

Penekanan relatif dari kedua prediksi dan deskripsi berbeda sehubungan dengan aplikasi yang mendasari dan teknik. Ada beberapa teknik data

mining memenuhi tujuan tersebut. Beberapa di antaranya adalah asosiasi, klasifikasi, dan clustering.

B. Clustering

Clustering adalah metode pengelompokan data ke dalam kelompok yang berbeda, sehingga data pada masing-masing kelompok memiliki kecenderungan dan pola yang sama. Clustering merupakan salah satu teknik dalam proses ekstraksi pengetahuan dalam data mining. Teknik ini secara otomatis akan mempartisi ruang data menjadi cluster atau kelompok. [6]. Gambar 1 berikut ini memperlihatkan beberapa algoritma yang ada pada teknik clustering:



Gambar 1. Teknik-teknik Clustering [7]

Clustering merupakan salah satu metoda yang paling penting di bidang data mining. Secara sederhana, clustering adalah cara untuk membagi dari data ke dalam kelompok yang berbeda. Data dikelompokkan ke dalam cluster dengan cara sedemikian rupa sehingga data dari kelompok yang sama memiliki tingkat kesamaan yang tinggi dan data dalam kelompok-kelompok lain memiliki perberbedaan. Hal ini bertujuan untuk meminimalkan kesamaan objek dalam kelas dan memaksimalkan perbedaan objek antar kelas. Clustering adalah teknik *unsupervised learning*. Dimana teknik ini berguna untuk menemukan pola yang menarik dari sekumpulan data yang besar. Clustering dapat diterapkan pada banyak area, seperti studi pemasaran, analisis DNA, perencanaan kota, penambangan teks, serta klasifikasi dokumen web dan lainnya yang berhubungan dengan pemartisian data.

C. Algoritma K-Medoids (Partitioning Around Medoids)

Algoritma PAM sering dikenal dengan algoritma k-Medoids adalah sebuah algoritma yang merepresentasikan cluster yang dibentuk menggunakan medoids. Proses pembentukan cluster dimulai dengan menentukan k objek dari

dataset secara acak sebagai medoid, selanjutnya hitung cost setiap objek non-medoid dengan k objek, cost terkecil setiap obyek non-Medoid terhadap medoid akan masuk dalam cluster dimana medoid tersebut berada. Secara iteratif ditentukan kembali objek baru secara acak, lalu proses perhitungan cost dilakukan kembali. Apabila total cost yang dihasilkan lebih kecil dari cost setiap objek dengan medoid lama, maka objek baru tersebut dapat menjadi medoid baru. Iterasi berakhir hingga tidak ada perubahan terhadap cost yang dihasilkan oleh medoid baru. Algoritma PAM, sebagai berikut [7]:

Input :

K: Jumlah cluster yang dikehendaki
D: Dataset yang berisi objek-objek

Output :

Cluster yang berisi objek-objek yang representatif didasarkan pada kedekatan minimum jarak dengan medoid

Metoda :

Secara acak, pilih k objek di dalam D sebagai medoid awal

Ulangi :

Tentukan objek terdekat dengan medoid

Secara acak pilih objek non-medoid O_{acak}

Hitung total point S untuk ditukarkan point O_j dengan O_{acak}

Jika $S < 0$ maka tukar O_j dengan O_{acak} sebagai k medoid baru

Hingga tidak ada perubahan

Untuk menghitung jarak antara objek-objek dengan medoids dapat menggunakan Euclidean Distance

$$\sqrt{\sum_{i=1}^k (x_i - y_i)^2} \quad (1)$$

Berdasarkan algoritma di atas, setiap objek dihitung jaraknya dengan medoids menggunakan perhitungan jarak Euclidean Distance pada Persamaan 1, jarak terkecil dari objek dalam data dikelompokkan dalam cluster tertentu sesuai dengan medoids-nya. X_i menggambarkan data pada medoids sedangkan y_i menggambarkan data pada obyek.

Untuk melihat mutu cluster dapat digunakan perhitungan Silhouette, yaitu sebuah teknik untuk melakukan evaluasi mutu dari sebuah cluster. Teknik ini merupakan metoda untuk melakukan validasi cluster, yaitu metoda yang menggabungkan metoda *cohesion* dan metoda *separation*. Tahapan yang dilakukan dalam metoda ini antara lain:

- Untuk setiap objek i, hitung rata-rata jarak dari obyek ke i dengan seluruh obyek yang berada pada satu cluster dan nilai rata-rata tersebut bernama a_i
- Untuk setiap objek i, hitung rata-rata jarak dari objek ke i dengan objek yang berada di cluster lainnya dan diambil nilai rata-rata terkecil, yang disebut b_i
- Hitung nilai koefisien Silhouette dengan Persamaan (2).

$$S_i = (b_i - a_i) / \max(a_i, b_i) \quad (2)$$

Hasil dari Silhouette dapat dikategorikan sebagai berikut :

- $0.7 < S_i \leq 1$ disebut sebagai cluster yang kuat (sangat baik)
- $0.5 < S_i \leq 0.7$ disebut sebagai cluster dengan struktur medium
- $0.25 < S_i \leq 0.5$ disebut sebagai cluster yang buruk
- $S_i \leq 0.25$ disebut cluster yang tidak memiliki struktur.

III. HASIL DAN DISKUSI

Algoritma PAM sering dikenal dengan algoritma k-Medoids adalah sebuah algoritma yang merepresentasikan cluster yang dibentuk menggunakan medoids. Proses pembentukan cluster dimulai dengan menentukan k objek dari dataset secara acak sebagai medoid, selanjutnya hitung cost setiap objek non-medoid dengan k objek, cost terkecil setiap objek non-medoid terhadap medoid akan masuk dalam cluster dimana medoid tersebut berada. Secara iteratif ditentukan kembali objek baru secara acak, lalu proses perhitungan cost dilakukan kembali. Apabila total cost yang dihasilkan lebih kecil dari cost setiap objek dengan medoid lama, maka objek baru tersebut dapat menjadi medoid baru. Iterasi berakhir hingga tidak ada perubahan terhadap cost yang dihasilkan oleh medoid baru. Algoritma PAM

- Pilih n objek secara acak sebagai medoid, dimana banyaknya n menggambarkan jumlah cluster yang diinginkan
- Hitung jarak terdekat setiap objek diluar medoid dengan masing-masing medoids, dari proses ini akan dipilih objek yang terdekat dengan medoidnya, sehingga kumpulan objek-objek tersebut akan berbentuk cluster
- Pilih secara acak objek pada masing-masing cluster sebagai kandidat medoid baru
- Hitung jarak setiap objek yang berada pada cluster masing-masing

dengan kandidat medoid, jika total jarak (cost) setiap objek dengan kandidat medoid < total jarak (cost) setiap objek dengan medoid, maka kandidat medoid menjadi medoid baru
5 ulangi proses 2 - 4 hingga tidak ada perubahan pada medoid

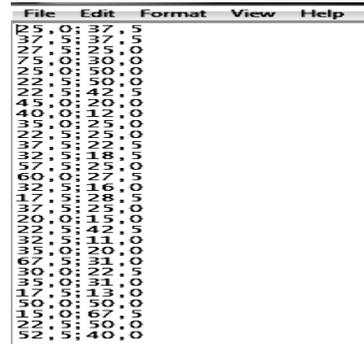
A. Data Percobaan

Pada penelitian ini akan digunakan data nilai mahasiswa yang mengambil mata kuliah algoritma pemrograman sebanyak 30 data dan data nilai mata kuliah struktur data sebanyak 30 orang, sehingga total data yang akan digunakan sebanyak 120 data. Adapun komponen nilai yang akan diolah dengan atribut : nilai Ujian Tengah Semester (UTS) dan Ujian Akhir Semester (UAS) mata kuliah Algoritma Pemrograman dan nilai UTS dan UAS mata kuliah Struktur Data. Ke 60 data yang digunakan masing-masing 30 data merupakan data mahasiswa Informatika Unjani yang mengikuti perkuliahan pada Tahun Akademik 2013/2014. Dari data tersebut akan dilakukan proses kalkulasi untuk menentukan cluster yang akan menggambarkan kemampuan mahasiswa dalam mengikuti kuliah Algoritma Pemrograman dan kuliah Struktur Data. Gambar 1 berikut ini adalah tabel nilai yang akan digunakan sebagai data percobaan. Data nilai ditulis menggunakan format spread sheet MS Excel.

	A	B	C
1	Struk.Data	Algoritma	
2	25	37,5	
3	37,5	37,5	
4	27,5	25	
5	75	30	
6	25	50	
7	22,5	50	
8	22,5	42,5	
9	45	20	
10	40	12	
11	35	25	
12	22,5	25	
13	37,5	22,5	
14	32,5	18,5	
15	57,5	25	
16	60	27,5	
17	32,5	16	
18	17,5	28,5	
19	37,5	25	
20	20	15	
21	22,5	42,5	
22	32,5	11	
23	35	20	
24	67,5	31	

Gambar 2. Data Percobaan Berupa Nilai Mata Kuliah Algoritma dan Struktur Data

Data nilai untuk mata kuliah algoritma dan struktur data terdiri dari nilai UTS dan nilai UAS yang dibuat secara rata-rata. Dataset di atas setelah dirata-ratakan berjumlah 30 data. Berdasarkan data tersebut proses selanjutnya yaitu memindahkan data dalam format MS Excell ke data text dengan delimiter (file csv). Gambar 2 memperlihatkan dataset yang telah diubah menjadi format comma delimited (CSV) sebagai berikut [1]:



B. Data nilai mahasiswa dengan format CSV
A. Penerapan Algoritma K-Medoids

Seperti pada pembahasan sebelumnya, bahwa algoritma PAM akan dimulai pada penentuan awal medoids dengan memilih secara acak di antara objek yang ada pada dataset. Jumlah objek yang dipilih sesuai dengan jumlah cluster yang diinginkan, hal ini dikarenakan objek yang terpilih akan merepresentasikan pusat dari cluster sebagai medoids, untuk kasus ini akan dipilih tiga buah objek sebagai medoid awal. Secara teknis dataset yang tersedia harus dibaca dan dipindahkan ke struktur larik dua dimensi, dimana setiap objek data akan bersesuaian dengan nomor indeks pada larik tersebut. Identifikasi objek dapat didasarkan pada nomor indeksinya. Tabel 1 struktur data larik dua dimensi yang akan menyimpan objek data yang dibaca

TABEL I. STRUKTUR DATA LARIK SEBAGAI DATASET

indeks	Uts1	Uas1	Uts2	Uas2
1	30	20	55	20
2	30	45	30	45
3	30	25	15	35
4	100	50	25	35
5	30	20	55	45
26		35	15	
27	30	30	20	25
28	20	5	15	11
29	30	70	70	30
30	30	0	85	50

Berdasarkan algoritma Tabel I, proses pembentukan cluster secara teknis akan digambarkan sebagai berikut :

1. Dipilih secara acak tiga buah objek dari dataset yang akan mewakili tiga buah cluster. Tiga buah objek tersebut akan merepresentasikan medoids (pusat cluster). Hasil dari pembentukan cluster awal seperti pada Gambar 3.

Penerapan Algoritma Partitioning Around Medoids (PAM)

Dalam Melihat Kemampuan Akademik Mahasiswa

Dibuat Oleh : Yulison Hery Christanto

Objek Sebagai Medoids Awal

Jumlah Data : 20

No. Urut Data	Nilai Struktur Data	Nilai Algoritma	Nomor Cluster
20	32.5	11.0	0
14	60.0	27.5	1
25	17.5	13.0	2

Gambar 3. Objek awal sebagai medoids

2. Tahapan berikutnya adalah menghitung jarak setiap objek non-medoids. Objek dengan jarak terdekat dengan masing-masing medoid akan bergabung dengan medoidnya. Gambar 4 berikut ini memperlihatkan pembentukan cluster dimulai dari iterasi ke-0 hingga iterasi ke-20. Sebagai catatan, jumlah iterasi yang digunakan sebanyak 20, hal ini terkait dengan konsumsi waktu yang diperlukan selama proses pembentukan cluster.

Pembentukan Cluster Berdasarkan Medoids Baru

Jumlah Iterasi : 100

Iterasi	Medoids Lama	Jarak Lama	Medoids Baru	Jarak Baru	Medoids Terpilih	Cluster
0	25	282.32415805136	5	170.68087441144	5	2
1	14	67.309077494984	26	118.75014016178	14	1
2	20	177.88407898999	9	139.62672251474	9	0
3	9	170.04053516623	15	175.6443822421	9	0
4	14	67.309077494984	22	70.774372538788	14	1
5	9	170.04053516623	11	172.5709270824	9	0
6	5	50.726440374977	19	56.336465611657	5	2
7	14	67.309077494984	13	75.72189502976	14	1
8	9	170.04053516623	10	221.93127336083	9	0
9	5	50.726440374977	28	50.726440374977	5	2
10	5	50.726440374977	27	140.03843208969	5	2
11	9	170.04053516623	2	181.83513131483	9	0
12	9	170.04053516623	2	181.83513131483	9	0
13	9	170.04053516623	23	159.05478745535	23	0
14	5	50.726440374977	27	140.03843208969	5	2
15	23	159.05478745535	12	162.85774548507	23	0
16	23	159.05478745535	16	295.20809614647	23	0
17	23	159.05478745535	17	181.13580150372	23	0
18	14	67.309077494984	29	83.846492943006	14	1
19	14	67.309077494984	26	118.75014016178	14	1

Gambar 4. Cluster yang terbentuk dalam 20 iterasi

Perhitungan jarak antara objek dengan setiap medoids menggunakan metoda Euclidean distance, sesuai dengan Persamaan 1

Euclidean Distance merupakan metoda perhitungan jarak yang didasarkan pada ruang berdimensi terbatas yang bernilai riil[5]. Jarak euclidean dua buah objek merupakan panjang sisi miring dari sebuah segi tiga siku-siku. Proses setiap iterasi berlaku algoritma k-Medoids (Tabel 1) dari langkah pertama hingga langkah 5, yaitu sebanyak 20 iterasi.

3. Menghitung mutu setiap cluster dengan menggunakan pendekatan Silhouette, dimana proses kalkulasi mutu mengikuti persamaan 2.2, Dimana **b** adalah rata-rata jarak medoid dengan objek di luar cluster, **a** adalah jarak rata-rata antara medoid dengan objek di dalam cluster. Gambar 5 memperlihatkan hasil

perhitungan menggunakan Silhouette bagi ketiga cluster, sebagai berikut :

Uji Mutu Cluster Dengan Silhouette

Cluster	Rata-rata Jarak Dalam Cluster	Rata-rata Jarak Luar Cluster	Silhouette
0	12.867484569159	31.743551883885	0.59464257130936
1	14.320509159578	36.40720513566	0.60665727824433
2	10.060068224419	33.09721956628	0.69604491385529

Gambar 5. Hasil uji mutu cluster dengan Teknik Silhouette

Berdasarkan hasil uji mutu cluster dalam Persamaan 2, nampak bahwa nilai Silhouette dari ketiga cluster mendekati angka 1, itu artinya cluster yang dibentuk sudah baik.

4. Melakukan analisa clustering untuk menginterpretasikan hasil dari cluster. Gambar 6 menunjukkan hasil rekapitulasi nilai akademik setiap clusternya, sebagai berikut :

Cluster Akhir yang Terbentuk

Nomor	Cluster	Nilai Struktur Data	Nilai Algoritma	Medoids
1	0	37.5	37.5	
2	0	27.5	25.0	
7	0	45.0	20.0	
8	0	40.0	12.0	
9	0	35.0	25.0	
10	0	22.5	25.0	
11	0	37.5	22.5	
12	0	32.5	18.5	
15	0	32.5	16.0	
16	0	17.5	28.5	
17	0	37.5	25.0	
18	0	20.0	15.0	
20	0	32.5	11.0	
21	0	35.0	20.0	
23	0	30.0	22.5	23
24	0	35.0	31.0	
25	0	17.5	13.0	
3	1	75.0	30.0	
13	1	57.5	25.0	
14	1	60.0	27.5	14
22	1	67.5	31.0	
26	1	50.0	50.0	
0	2	25.0	37.5	
4	2	25.0	50.0	
5	2	22.5	50.0	5
6	2	22.5	42.5	
19	2	22.5	42.5	
27	2	15.0	67.5	

Gambar 6. Objek data dalam setiap cluster

Berdasarkan Gambar 6, maka dapat diinterpretasikan hasil kemampuan akademik mahasiswa khususnya pada mata kuliah algoritma dan struktur data sebagai berikut : Cluster-0 menyatakan kumpulan nilai mahasiswa secara umum rendah (56,66 %), cluster-1 menyatakan kumpulan nilai mahasiswa secara sedang (23,3 %), dan Cluster-2, menyatakan kumpulan nilai mahasiswa baik (16,66%). Interpretasi dari hasil clustering di atas seperti pada Gambar 7.

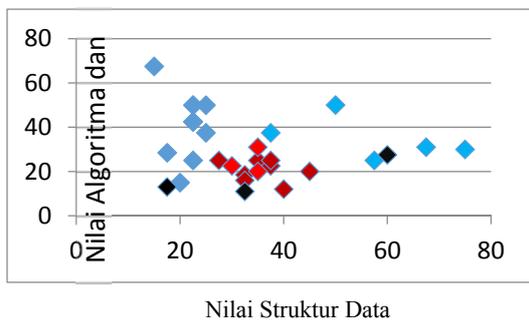
Analisa Clustering/ Interpretasi

Cluster	Jumlah Data	Nilai Rata-rata Struktur Data	Nilai Rata-rata Algoritma
0	17	31.18	21.47
1	5	61.80	32.60
2	7	21.86	48.29

Gambar 7. Hasil interpretasi dari clustering

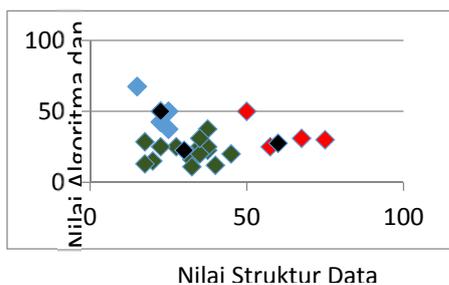
Dalam proses clustering menggunakan algoritma K-Medoids didapatkan hasil cluster berdasarkan medoid yang tidak berubah lagi setelah proses penentuan medoid pada iterasi ke-14. Berikut ini adalah cluster yang terbentuk pada awal iterasi hingga akhir iterasi, yaitu sebanyak 20 iterasi sebagai berikut.

Pada Gambar 8 berikut ini merupakan cluster awal yang dibentuk berdasarkan medoid yang dibangkitkan secara acak sebanyak 3 (tiga) medoid. Setiap objek non-medoid merupakan anggota cluster tertentu, dimana objek tersebut memiliki jarak terdekat dengan masing-masing medoid.



Gambar 8. Pembentukan cluster awal

Sedangkan pada Gambar 9 merupakan cluster akhir yang terbentuk, yang didasarkan pada perubahan medoid hingga iterasi ke-14, dimana tidak terjadi lagi perubahan medoid.



Gambar 9. Pembentukan cluster akhir

IV. KESIMPULAN

Partitioning around medoids (PAM) atau lebih dikenal dengan algoritma k-Medoids merupakan salah satu algoritma clustering data mining. Dengan data uji berupa data nilai mahasiswa yang mengikuti mata kuliah algoritma pemrograman dan struktur data masing-masing sebanyak 30 data, dihasilkan cluster yang merepresentasikan keadaan umum kemampuan akademik mahasiswa khususnya pada mata kuliah tersebut. Algoritma k-Medoids telah bekerja dengan baik untuk

menentukan setiap objek pada dataset untuk setiap cluster dengan jumlah iterasi sebanyak 20 dengan jumlah cluster sebanyak 3. Mutu setiap cluster telah diuji menggunakan teknik Silhouette, dan hasil dari pengujian untuk setiap cluster memiliki nilai Silhouette mendekati 1, dengan demikian setiap cluster dapat dikatakan memiliki mutu yang baik, dimana setiap objek telah dikelompokkan sesuai dengan tingkat kesamaan yang tinggi. Untuk penulis lebih dua institusi, ubah dengan langkah.

UCAPAN TERIMAKASIH

Ucapan terimakasih disampaikan kepada LPPM Unjani yang telah memberikan kesempatan yang luas untuk penelitian ini sehingga penelitian ini dapat dilakukan sesuai dengan harapan. Ucapan terimakasih juga disampaikan kepada Jurusan Informatika Unjani atas data akademik yang digunakan sebagai bahan uji coba, khususnya untuk mata kuliah algoritma pemrograman dan struktur data.

DAFTAR PUSTAKA

- [1] Chrisnanto, Yulison Herry, 2013, Algoritma Partitioning Around Medoids (PAM) Sebagai Teknik Clustering Pada Data Mining, Majalah Ilmiah Aristoteles FMIPA Unjani, Volume 11 Nomor 2, Oktober 2013.
- [2] Grabust, Peter. The choice of metrics for clustering algorithms. Proceeding of the 8th International Scientific and Practical Conference. Volume II, 2011, ISSN 1691-5402, ISBN 978-9984-44-071-2
- [3] Han, Jiawei and Kamber, Micheline, 2006, Data Mining: Concept and Techniques, second editions, Elsevier, Morgan Kaufman Publisher.
- [4] Ibrahim, Lamiaa Fattouh, Al Harbi, Manal Mamed. Using Modified Partitioning Around Medoids Clustering Technique in Mobile Network Planning. <http://arxiv.org/ftp/arxiv/papers/1302/1302.6602.pdf>
- [5] Kumari, Archana, Bhagat, Hritu. Compression Record Based Efficient k-Medoid Algorithm Increase Scalability dan Efficiency. International Journal of Advanced Research in Computer Engineering and Technology (IJARCET), volume 2 issue 8, August 2013.
- [6] Wibisono, Yudi, 2011, Perbandingan Partitioning Around Medoids (PAM) dan k-Means Clustering Untuk Tweets, KNSI2011, Februari 2011.