

CLUSTERING PROGRAM KEAHLIAN PADA PENDAFTARAN SISWA BARU (PSB) DENGAN MENGGUNAKAN ALGORITMA K-MEANS: STUDI KASUS DI SMKN 1 NGULING

Endang Widia Astuti¹

¹) Kompetensi Keahlian Multimedia SMK Negeri 1 Nguling

sttastuti@gmail.com¹

Abstrak

Dalam dunia pendidikan, khususnya di lembaga pendidikan SMK (Sekolah Menengah Kejuruan) rata-rata sekolah memiliki beberapa (lebih dari satu) program keahlian sehingga memungkinkan terjadi kebingungan bagi siswa yang ingin mendaftar dalam memilih jurusan yang sesuai dengan keinginan. Tentunya dalam mendaftar seorang siswa akan memiliki acuan/tolak ukur untuk memilih jurusan yang sesuai dengan bakat, minat dan keinginannya. Hal tersebut tentunya perlu adanya informasi yang tepat sehingga siswa benar-benar paham tentang jurusan yang akan dipilihnya.

Oleh karena itu perlu suatu Data Mining untuk membantu mengambil keputusan siswa dalam memilih jurusan sehingga mempermudah siswa dalam mengambil keputusan secara tepat dan akurat. Penelitian ini akan menggunakan algoritma K-Means, karena dalam algoritma K-Means acuannya adalah angka yang dikelompokkan/ clustering diharapkan siswa langsung mengetahui jurusan yang tepat yang sesuai dengan keputusan dari Data Mining.

Kata kunci : SMK, data mining, classification model, K-Means, Clustering.

1. PENDAHULUAN

Dunia pendidikan merupakan salah satu aspek yang menentukan dalam menghasilkan penerus bangsa. Dalam hal ini dari pengetahuan khususnya dilingkungan tingkat SMK, ada beberapa jurusan di SMK yang akan dipilih oleh siswa yang akan masuk ke sekolah tersebut. Tentunya siswa dalam memilih jurusan di SMK tersebut harus disesuaikan dengan skill/ keahlian yang dimiliki dan didukung oleh informasi yang tepat baik berupa informasi lisan dari alumni atau masyarakat sekitar dan informasi yang lainnya agar siswa tidak salah dalam memilih jurusan.

Adapun jurusan yang ada di SMK antara lain

- A. Multimedia,
- B. Teknologi Hasil Pertanian Pangan, dan
- C. Teknologi Kendaraan Ringan

Dengan Data Mining yang diolah berdasarkan informasi yang ada, nantinya siswa akan lebih tepat dalam mengambil jurusan di SMK sehingga dapat mempermudah dalam menentukan masa depannya.

2. LANDASAN TEORI

A. Data, Informasi dan Database

Sebelum kita membahas tentang data mining, hal yang harus dipahami terlebih dahulu yaitu pengertian tentang data, sistem informasi dan database.

1. Data

Data merupakan fakta, gambar atau suara yang mungkin atau tidak berhubungan atau berguna bagi tugas tertentu. Data terdiri dari fakta-fakta dan angka yang secara relatif tidak berarti bagi pemakai [McLeod 2004]

Data merupakan suatu bentuk keterangan-keterangan yang belum diolah atau dimanipulasi sehingga belum begitu berarti bagi sebagian pemakai. Data yang kita bicarakan dalam konteks data mining yang sebenarnya adalah data yang dikelola dalam suatu database container. Data yang dialokasikan pada database demikian dikelola oleh database server khusus dan memang digunakan oleh perusahaan berskala menengah ke atas. Informasi yang terkandung didalamnya berisi jutaan record dan selalu mengalami modifikasi seraya perusahaan bertransaksi dan beraktifitas.

2. Informasi

Informasi adalah data yang sudah diolah sehingga menjadi lebih berguna bagi pemakai. Informasi dapat diibaratkan sebagai darah yang mengalir di dalam tubuh manusia, seperti halnya informasi di dalam sebuah perusahaan yang sangat penting untuk mendukung kelangsungan perkembangannya, sehingga terdapat alasan bahwa informasi sangat dibutuhkan bagi sebuah perusahaan. Akibat kurangnya informasi dalam waktu tertentu perusahaan akan mengalami ketidak mampuan mengontrol sumber daya, sehingga dalam mengambil keputusan strategis sangat terganggu.

3. Basis Data (data base)

Basis data merupakan sekelompok tabel data berisi informasi yang saling berelasi atau berhubungan. Data didalam database diproses untuk mengidentifikasi berbagai alternatif dan mengevaluasi potensinya untuk memecahkan masalah, setelah teridentifikasi solusi terbaik dilaporkan ke organisasi. Database bahkan dapat menyimpan solusi terbaik untuk suatu masalah tertentu sehingga bila dimasa depan organisasi menghadapi masalah yang sama, tersedia suatu solusi tanpa perlu memproses ulang data. Database atau basis data dapat di golongan berdasarkan ragam criteria, salah satunya berdasarkan frekuensi perubahannya (seberapa sering database mengalami perubahan) di kenal dua kategori yaitu :

1. Data master (data yang jarang berubah , misalnya data nama siswa, dan nomor induk siswa).
2. Data Transaksi atau data operasional (data yang sering berubah dan membutuhkan data master untuk membangun komponennya. Misalnya data siswa, guru dan lain-lain).

Data master jarang berubah, karena data siswa seperti nama dan nomor induk siswa, bahkan tidak mengalami perubahan dalam jangka waktu yang lama. Tapi bukan berarti pula data tersebut tidak berubah. Misalnya alamat siswa berubah otomatis data master siswa ikut berubah. Sedangkan data transaksi akan selalu berubah karena aktivitas sekolah, misalnya siswa yang mutasi.

B. Data Mining

1. Definisi data Mining

Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Data mining adalah proses yang menggunakan teknik statistic, matematika, kecerdasan buatan, dan machine e-learning untuk mengekstrasi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar [Turban 2005].

Menurut gartner group, data mining adalah suatu proses menemukan hubungan yang berarti, pola dan kecenderungan dengan memeriksa dan sekumpulan besar data yang tersimpan dalam penyimpanan dengan menggunakan teknik pengenalan pola seperti teknik statistic dan matematika [Larose 2005].

2. Komponen Data Mining

Secara alami, material data mining sebenarnya sudah terbentuk karena factor rutinitas dan waktu seraya perusahaan melakukan aktivitasnya. Tanpa disadari perusahaan berinvestasi dengan menggunakan budgetnya untuk penggunaan teknologi informasi atau computer. Teknologi data mining mulai muncul karena akumulasi data yang besar dan pesat pertumbuhannya sehingga menimbulkan apa yang disebut rich of data but poor information. Tumpukan data ini tidak dapat digunakan pada aplikasi yang ada sehingga menjadi gudang data.

3. Pengelompokan data mining

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan , yaitu [larose, 2005]

1. Deskripsi

Terkadang peneliti dan analis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data.

2. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variable target. Estimasi lebih kearah numeric dari pada kearah kategori.

3. Prediksi

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada dimasa datang.

4. Klasifikasi

Dalam klasifikasi terdapat target variable kategori, sebagai contoh penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah.

5. Pengklusteran

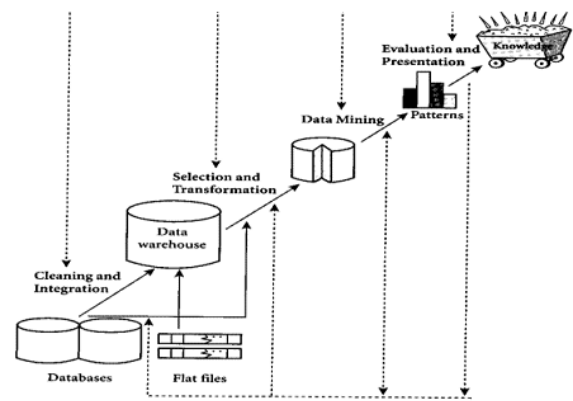
Pengklusteran merupakan pengelompokan record, pengamatan atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan. Kluster adalah kumpulan record yang memiliki kemiripan satu dengan yang lainya dan memiliki ketidak miripan dengan record-record dalam cluster.

6. Asosiasi

Tugas asosiasi dalam data mining adalah menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang biasa.

4. Tahapan Data Mining

Tahapan data mining dilakukan melalui pembentukan data warehouse karena hanya struktur data dari data warehouse yang dapat digunakan untuk data mining. Dengan melaksanakan ekstraksi, pembersihan dan transformasi (Extraction, Cleansing, Transformation) atas data transaksional yang disebut dengan data OnLine Transactional Processing (OLTP) ke OnLine Analytical Processing (OLAP) System untuk membuat sebuah data warehouse maka akan didapat suatu kumpulan data yang luas dan besar serta bersifat statis dan historis yang berasal dari data-data transaksi OLTP selama perusahaan menjalankan bisnisnya [Vieira 2000].



Gambar 2.1 Tahapan Data Mining [Pramudiono 2003]

5. Algoritma Data Mining

K-means merupakan metode clustering secara partitioning yang memisahkan data ke dalam kelompok yang berbeda dengan partitioning secara iteratif, K-Means mampu meminimalkan rata-rata jarak setiap data ke cluster-nya.

Algoritma K-means adalah sebagai berikut :

- a. Tentukan nilai k sebagai jumlah klaster yang ingin dibentuk.
- b. Menentukan nilai centroid

Dalam menentukan nilai centroid (titik pusat klaster) untuk awal iterasi nilai awal dilakukan secara random. Sedangkan jika menentukan nilai centroid yang merupakan tahap dari iterasi, maka digunakan rumus sebagai berikut.

$$\bar{v}_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} x_{kj} ,$$

dimana :

- \bar{v}_{ij} adalah centroid/ rata-rata cluster ke-I untuk variable ke-j.
- N_i adalah jumlah data yang menjadi anggota cluster ke-i
- i,k adalah indeks dari cluster
- j adalah indeks dari variabel
- x_{kj} adalah nilai data ke-k yang ada di dalam cluster tersebut untuk variable ke-j

- c. Menghitung jarak antara centroid dengan titik tiap objek, untuk menghitung jarak tersebut dapat menggunakan rumus korelasi antar dua objek yaitu *Euclidean Distance*

$$D_e = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2}$$

dimana :

D_e adalah Euclidean Distance

i adalah banyaknya objek,

(x,y) merupakan koordinat object dan

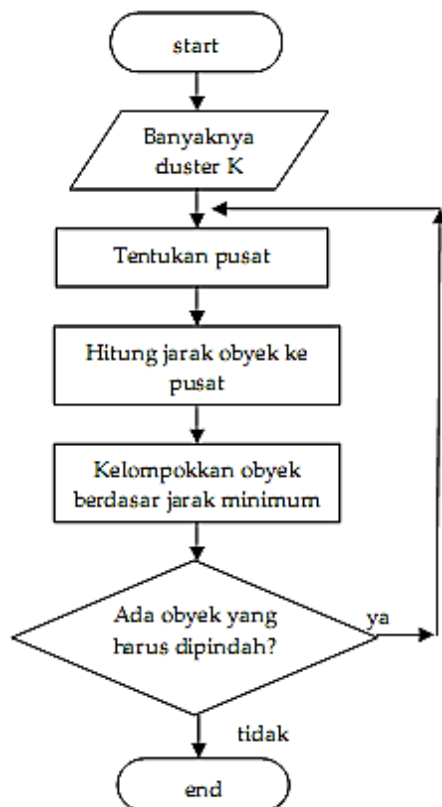
(s,t) merupakan koordinat centroid.

d. Pengelompokan Objek

Untuk menentukan anggota cluster adalah dengan memperhitungkan jarak minimum objek. Nilai yang diperoleh dalam keanggotaan data pada distance matriks adalah 0 atau 1, dimana nilai 1 untuk data yang dialokasikan ke cluster dan nilai 0 untuk data yang dialokasikan ke cluster yang lain.

- e. Kembali ke tahap 2, lakukan perulangan hingga nilai centroid yang dihasilkan tetap dan anggota cluster tidak berpindah ke cluster lain.

Flowchart K-Means Clustering



Gambar 1. Flowchart algoritma K-Means

3. TINJUAN OBJEK PENELITIAN

Sejarah SMKN 1 Nguling

SMKN1 NGULING didirikan berdasarkan surat keputusan bupati pasuruan no. 1633/579/HK/424.022/2003 tanggal 18 – 11 – 2003. NSS.32.1.05.19.21.020 di bangun diatas lahan SMPN 1 NGULING di desa Nguling (jl. Dr. Soetomo no. 69 Nguling dengan 5.109 m². SMKN 1 NGULING didirikan dari anggaran pemerintah pusat program SMK di SMP (SMK terpadu pada waktu itu) dan terus berkembang sehingga sekarang menjadi SMKN 1 Nguling. Bersamaan dengan itu pemerintah kab. Pasuruan membuka sekolah menengah kejuruan negeri yaitu smkn prigen, smkn tutur, smkn puspo, smkn winongan dan smkn 1 nguling. Pembangunan pertama adalah 2 RKB: 1 bengkel, 1 lab Komputer dan 1 kantor dengan jumlah program keahlian pada awal pendirian adalah teknik mekanik otomotif (TMO) sekarang menjadi teknik kendaraan ringan (TKR) dan program keahlian teknologi hasil pertanian (THP) sekarang menjadi (TPHP).

PERKEMBANGAN PROGRAM KEAHLIAN

Jumlah peserta didik terdiri dari 3 rombel: 2 rombel teknik mekanik otomotif = 50 siswa, 1 rombel THP = 6 siswa. Pada tahun 2009 – 2010 dibuka program keahlian multimedia, mengingat banyak animo masyarakat yang menghendaki program multimedia. Alhamdulillah tahun 2015 – 2016 ini jumlah program keahlian 3 program. Jumlah rombel 22 rombel;

- A. TKR = 9 Rombel
- B. MM = 7 Rombel
- C. TPHP = 6 Rombel

Jumlah peserta didik = 770 siswa

4. DESAIN PENELITIAN

1. Metode Penelitian

Agar penelitian yang dilakukan lebih terarah maka peneliti menggunakan suatu metode kuantitatif yang bertujuan untuk mendapatkan angka yang dapat diperoleh dari proses dan subjek penelitian.

2. Metode Pengumpulan Data

Metode dalam pengumpulan data yang adalah menggunakan kuisisioner yang diberikan

kepada calon Siswa baru dan berdasarkan nilai yang diperoleh pada UN.

3. Teknik Analisa Data

Data yang telah diperoleh dari Kuisisioner yang berupa angka kuantitatif nantinya diproses dengan algoritma K-Means.

4. Instrument

Instrumen yang digunakan peneliti untuk menunjang proses penelitian adalah menggunakan Microsoft excel 2010, Microsoft office 2010, Xampp versi 17.4.

5. ANALISIS, INTERPRETASI, DAN IMPLIKASI PENELITIAN

A. Persiapan Data

Dalam memperoleh data kita menggunakan Kuisisioner berupa angket yang berisi data untuk memilih, dimana dalam angket tersebut terdapat nilai-nilai berupa angka yang nantinya dijadikan data untuk proses analisa.

B. Teknik Data Mining

Teknik yang Kami gunakan untuk pengelompokkan Jurusan berdasarkan 3 kriteria, yaitu kemampuan(*skill*) dasar, bakat, dan minat adalah dengan menggunakan algoritma K-Means, yang dimulai dengan pembentukan cluster di awal kemudian secara iteratif prototipe cluster ini diperbaiki hingga konvergen (tidak terjadi perubahan yang signifikan pada prototipe cluster). Perubahan ini diukur menggunakan jumlah atau rata-rata(*means*) jarak tiap item data dengan Objek.

C. Testing dan Implementasi

Data yang diperoleh dari hasil kuisioner adalah sebagai berikut :

N o	Nama	Skill Dasar	Bakat	Minat
1	A. MUKHLASON	30	25	45
2	A. HABIBULLAH	30	40	30
3	A. SIDDIQ R	20	60	20
4	ARDIANTO	40	30	30
5	AYU SETYO	70	20	10
6	BAHRI P	10	20	10

1. Menentukan Jumlah Cluster

Dengan memperoleh data tersebut, maka kita kelompokkan objek tersebut kedalam 3 cluster dengan atribut TPHP, TKR, MM, seperti tabel dibawah ini :

N o	Skill Dasar	Bakat	Minat	TPHP	TKR	MM
1	30	25	45			
2	20	70	10			
3	20	60	20			
4	40	30	30			
5	70	20	10			
6	10	20	10			

2. Menentukan Nilai Centroid

Untuk menentukan nilai centroid ditentukan secara acak, kita mengambil nilai centroid :

- TPHP = (25,20,25)
- TKR = (20,25,30)
- MM = (30,25,25)

3. Menghitung jarak antara titik centroid dengan tiap titik objek

Untuk menghitung jarak antara titik centroid dengan tiap titik menggunakan rumus Euclidean Distance, yaitu :

$$D_e = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2}$$

dimana :

D_e adalah Euclidean Distance

i adalah banyaknya objek,

(x,y) merupakan koordinat object dan

(s,t) merupakan koordinat centroid.

Cara menghitung :

N o	Skill Dasar	Bakat	Minat	TPHP			TKR			MM		
				2	2	2	2	2	3	3	2	2
1	30	25	45									
2	30	40	30									
3	20	60	20									
4	40	30	30									
5	70	20	10									
6	10	20	10									

Calon siswa 1(30,25,45) dengan TPHP (25,20,25)

$$\rightarrow = \sqrt{(30 - 25)^2 + (25 - 20)^2 + (45 - 25)^2} = 21.21320344$$

Dengan TKR (20,25,30)

$$\rightarrow = \sqrt{(30 - 20)^2 - (25 - 20)^2 - (45 - 25)^2} = 18.02775638 = 18.03$$

Dengan MM = (30,25,25)

$$\rightarrow = \sqrt{(30 - 30)^2 - (25 - 25)^2 - (45 - 25)^2} = 20$$

Dari perhitungan diatas dapat diperoleh tabel dibawah ini :

TPHP			TKR			MM		
Centroid 1			Centroid 2			Centroid 3		
25	20	25	20	25	30	30	25	25
40.62019			42.72002			36.05551		
21.2132			18.02776			15.81139		
40.62019			36.40055			36.74235		
18.70829			20.61553			12.24745		
47.43416			54.08327			43.01163		
21.2132			22.91288			25.4951		

Dengan distance matriknya seperti dibawah ini:

$$D^0 = \begin{matrix} 21.21 & 21.21 & dst & \longrightarrow & TPHP \\ 18.03 & 18.03 & dst & \longrightarrow & TKR \\ 20 & 15.81 & dst & \longrightarrow & MM \end{matrix}$$

Siswa1 Siswa2 SiswaX

4. Pengelompokan Objek

Setelah menghitung distance matriks, kita menentukan anggota cluster menurut jarak minimum dari centroid. Dengan merujuk pada distance matriks. Hal ini dapat dilihat pada perolehan nilai sebagai berikut :

			TPHP			TKR			MM					
			Centroid 1			Centroid 2			Centroid 3					
Skill	Bakat	Minat	25	20	25	20	25	30	30	25	25	Centroid 1	Centroid 2	Centroid 3
60	25	45	40.62019			42.72002			36.05551					ok
30	40	30	21.2132			18.02776			15.81139					ok
20	60	20	40.62019			36.40055			36.74235				ok	
40	30	30	18.70829			20.61553			12.24745					ok
70	20	10	47.43416			54.08327			43.01163					ok
10	20	10	21.2132			22.91288			25.4951			ok		
30	50	20	30.82207			28.72281			25.4951					ok

5. Iterasi 1, menentukan centroid baru

Himpunan yang terbentuk pada tahap sebelumnya, telah diketahui anggota tiap cluster. Dari data tersebut, hitungkembali centroid untuk menentukan centroid baru. Karena pada cluster 1 hanya mempunyai 1 anggota, maka untuk centroid baru masih berada di C1 = (10,20,10). Sedangkan pada C2& C3 dengan menghitung nilai rata-

ratanya dapat diperoleh nilai centroid barunya, yaitu :

$$C2 = ((30+20)/2, (25+60)/2, (45+20)/2) = (25, 42.5, 32.5)$$

$$C3 = ((30+40+70)/3, (40+30+20)/3, (30+30+10)/3) = (46.7, 30, 23.33)$$

Digambarkan seperti dibawah ini :

TPHP			TKR			MM		
Centroid 1			Centroid 2			Centroid 3		
25	20	25	20	25	30	30	25	25
10	20	10	40	42.5	32.5	46.7	30	23.33

6. Iterasi 1, menghitung jarak antara titik centroid baru dengan titik objek.

Pada langkah ke-6 ini merupakan pengulangan pada proses ke-2, sehingga hasil yang didapat seperti dibawah ini:

TPHP1			TKR1			MM1		
Centroid 1			Centroid 2			Centroid 3		
10	20	10	40	42.5	32.5	46.7	30	23.33
61.24			29.37			25.93		
34.64			10.61			20.55		
42.43			29.37			40.28		
37.42			12.75			9.43		
60.00			43.73			28.67		
0.00			43.73			40.28		
37.42			17.68			26.25		
24.49			66.80			60.18		

7. Iterasi 1, Pengelompokan Objek

Hampir sama dengan tahap 4, yaitu menentukan anggota cluster dengan menghitung jarak minimum tiap object dengan centroid baru. Hasil yang diperoleh :

TPHP1			TKR1			MM1					
Centroid 1			Centroid 2			Centroid 3					
10	20	10	40	42.5	32.5	46.7	30	23.33	Centroid 1	Centroid 2	Centroid 3
61.24			29.37			25.93					ok
34.64			10.61			20.55				ok	
42.43			29.37			40.28				ok	
37.42			12.75			9.43					ok
60.00			43.73			28.67					ok
0.00			43.73			40.28			ok		
37.42			17.68			26.25				ok	

8. Iterasi 2, Menentukan Centroid baru

Tahap ini mengulang kembali tahap 5, yaitu menghitung centroid baru. Hasil yang didapat adalah sebagai berikut :

TPHP1			TKR1			MM1		
Centroid 1			Centroid 2			Centroid 3		
10	20	10	40	42.5	32.5	46.7	30	23.33

9. Iterasi 2, menghitung jarak antara titik centroid baru dengan tiap titik object.

Tahap ini juga hampir sama dengan tahap 3, yaitu menghitung jarak dengan Centroid baru. Dengan cara perhitungan yang sama pada tahap 3, maka diperoleh hasil, yaitu

TPHP2			TKR2			MM2		
Centroid 1			Centroid 2			Centroid 3		
10	20	10	37	42	32	55	25	20
61.24			31.62			25.50		
34.64			7.07			30.82		
42.43			27.39			49.50		
37.42			12.25			18.71		
60.00			45.28			18.71		
0.00			40.62			46.37		

10. Iterasi 2, melakukan pengelompokan object

Hampir sama dengan tahap 4, yaitu menentukan anggota cluster dengan menghitung jarak minimum tiap object dengan centroid baru yang telah dihasilkan. Hasil yang diperoleh :

TPHP2			TKR2			MM2					
Centroid 1			Centroid 2			Centroid 3			Centroid 1	Centroid 2	Centroid 3
10	20	10	37	42	32	55	25	20			
61.24			31.62			25.50					ok
34.64			7.07			30.82				ok	
42.43			27.39			49.50				ok	
37.42			12.25			18.71				ok	
60.00			45.28			18.71					ok
0.00			40.62			46.37			ok		

11. Iterasi 3, Menentukan Centroid baru

Tahap ini mengulang kembali tahap 5, yaitu menghitung centroid baru. Hasil yang didapat adalah sebagai berikut :

Centroid 1			Centroid 2			Centroid 3		
10	20	10	50	51.7	41.7	70	20	10

12. Iterasi 3, menghitung jarak antara titik centroid baru dengan tiap titik object.

Tahap ini juga hampir sama dengan tahap 3, yaitu menghitung jarak dengan Centroid baru. Dengan cara perhitungan yang sama pada tahap 3, maka diperoleh hasil, yaitu

TPHP3			TKR3			MM3		
Centroid 1			Centroid 2			Centroid 3		
10	20	10	50	51.7	41.7	70	20	10
61.24			28.67			36.74		
34.64			25.93			48.99		
42.43			37.93			64.81		
37.42			26.56			37.42		
60.00			49.05			0.00		
0.00			60.05			60.00		

13. Iterasi 3, melakukan pengelompokan object

Hampir sama dengan tahap 4, yaitu menentukan anggota cluster dengan menghitung jarak minimum tiap object dengan centroid baru yang telah dihasilkan. Hasil yang diperoleh :

TPHP3			TKR3			MM3					
Centroid 1			Centroid 2			Centroid 3			Centroid 1	Centroid 2	Centroid 3
10	20	10	50	51.7	41.7	70	20	10			
61.24			28.67			36.74				ok	
34.64			25.93			48.99				ok	
42.43			37.93			64.81				ok	
37.42			26.56			37.42				ok	
60.00			49.05			0.00					ok
0.00			60.05			60.00			ok		
37.42			29.53			50.99				ok	

Berdasarkan hasil anggota cluster yang diperoleh tetap sama antara Iterasi2 = Iterasi3, maka iterasi dihentikan.

6. Penutup

A. Kesimpulan

Dari penelitian yang kami lakukan, dapat disimpulkan :

- 1) K-means Algoritma merupakan algoritma yang sederhana
- 2) K-means clustering mampu menyelesaikan permasalahan yang ada.
- 3) Hasil yang diperoleh dari penelitian diatas dapat disimpulkan pada tabel dibawah ini

NO	Nama	Hasil Kuisisioner			Jurusan Yang sesuai		
		Skill	Bakat	Minat	TPHP	TKR	MM
1	A. MUKHLASON	60	20	45		✓	
2	A. HABIBULLAH	30	40	30		✓	
3	A. SIDDIQ R	20	60	20		✓	
4	ARDIANTO	40	30	30		✓	
5	AYU SETYO	70	20	10			✓
6	BAHRI P	10	20	10	✓		

B. Saran

Mengingat lembaga SMKN 1 Nguling merupakan salah satu sekolah Rujukan di wilayah Nguling Kab Pasuruan dan memiliki banyak siswa, maka perlu adanya suatu penelitian lebih lanjut dalam menganalisis. Utamanya yang berhubungan dengan calon

siswa yang akan masuk ke SMKN 1 Nguling
Supaya siswa nantinya yang masuk benar-
benar sesuai dengan apa yang diharapkan baik
bagi siswa itu sendiri dan tujuan dari jurusan
serta sekolah itu sendiri.

7. Daftar Pustaka

- E.S. Gopi, "Algorithm Collections for Digital
Signal Processing Applications Using
Matlab", Spinger: National Institute of
Technology, Tiruchi, India,
Chen Yu, "K-Means Clustering", Indiana
University
Sergios Theodoridis, Konstantinos Koutroumbas
: "Pattern Recognition", Elsevier
Academic Press
Teknomo, Kardi. K-Means Clustering Tutorials.
[http://people.revoledu.com/kardi/tutorial/
kMean/](http://people.revoledu.com/kardi/tutorial/kMean/)
http://en.wikipedia.org/wiki/k-means_algorithm
[http://home.dei.polimi.it/matteucc/clustering/tuto
rial_ht](http://home.dei.polimi.it/matteucc/clustering/tutorial_ht)
Agusta, Y. (2004). Minimum Message Length
Mixture Modelling for Uncorrelated and
Correlated Continuous Data Applied to
Mutual Funds Classification, Ph.D.
Thesis, School of Computer Science and
Software Engineering, Monash
University, Clayton, 3800 Australia

