

PENERAPAN METODE *K-MEANS CLUSTER ANALYSIS* PADA SISTEM PENDUKUNG KEPUTUSAN PEMILIHAN KONSENTRASI UNTUK MAHASISWA *INTERNATIONAL CLASS* STMIK AMIKOM YOGYAKARTA

Josi Aranda¹⁾, Wirda Astari Galvani Natasya²⁾

^{1), 2)} Teknik Informatika STMIK AMIKOM Yogyakarta
Jl Ring road Utara, Condongcatur, Sleman, Yogyakarta 55281
Email : josi.a@students.amikom.ac.id¹⁾, wirda.n@students.amikom.ac.id²⁾

Abstrak

Penentuan pemilihan konsentrasi pada Perguruan Tinggi hampir sama halnya dengan pemilihan jurusan pada SMA. Penentuan pemilihan konsentrasi bukanlah hal yang mudah dan akan sangat berpengaruh terhadap kegiatan akademik mahasiswa. Keputusan penentuan konsentrasi harus dilakukan dengan pertimbangan yang matang. Terkadang ketika mahasiswa tersebut tidak mengetahui kemampuannya, mereka akan cenderung mengikuti trend pilihan teman-temannya. Hal ini tentu akan mempengaruhi kegiatan akademik mahasiswa tersebut, tidak dapat mengikuti perkuliahan dengan baik dan mungkin akan membuat mahasiswa tersebut menyesal dan merasa salah dalam menentukan pilihan konsentrasin perkuliahannya. Untuk membantu dalam memberikan dukungan dan kemudahan penentuan keputusan dalam memilih konsentrasi dibutuhkan sebuah Sistem Pendukung Keputusan Memilih Konsentrasi untuk Mahasiswa.

Pada penelitian ini akan digunakan metode *K-Means Cluster Analysis* untuk membantu dalam menentukan kecenderungan pemilihan konsentrasi studi mahasiswa dari nilai matakuliah yang ada. Dengan adanya penelitian ini diharapkan setiap mahasiswa dapat mengetahui dan fokus pada kemampuannya, dan kelak mampu mengikuti materi yang disampaikan juga sesuai dengan nilai yang dimilikinya. Adapun input yang digunakan pada Sistem Pendukung Keputusan menggunakan metode *K-Means Clustering* ini adalah Indeks Prestasi (IP) semester 3 matakuliah terkait konsentrasi yang ditawarkan.

Hasil dari penelitian ini adalah Pengujian yang dilakukan dalam penelitian ini, iterasi clustering data mahasiswa terjadi sebanyak 3 kali iterasi. Berdasarkan dari hasil Cluster jarak terdekat jika menunjukkan 1 maka diarahkan untuk konsentrasi Programming, 2 Multimedia dan 3 Jaringan Komputer. Setelah dilakukan pengolahan data dan mengalami 3 kali iterasi clustering maka ditemukan hasil sebagai berikut, 4 dari 12 mahasiswa diarahkan untuk mengambil konsentrasi Pemrograman dan 4 mahasiswa diarahkan untuk mengambil konsentrasi Multimedia sementara 3-5 mahasiswa diarahkan untuk mengambil konsentrasi Jaringan Komputer.

Kata kunci: Pemilihan Konsentrasi, *K-Means Clustering*, Sistem Pendukung Keputusan, Akademik Mahasiswa

1. Pendahuluan

Penentuan pemilihan konsentrasi pada Perguruan Tinggi sama halnya dengan pemilihan jurusan pada SMA. Penentuan pemilihan konsentrasi bukanlah hal yang mudah dan akan sangat berpengaruh terhadap kegiatan akademik mahasiswa. Keputusan penentuan konsentrasi harus dilakukan dengan pertimbangan yang matang. Banyak mahasiswa yang masih belum mengetahui minat, bakat dan kemampuan yang dimilikinya. Mereka akan cenderung mengikuti *trend* pilihan teman-temannya dan membuat mereka akan menjalani konsentrasi yang tidak sesuai dengan kemampuannya.

Banyaknya pilihan konsentrasi yang ditawarkan juga menjadi alasan mahasiswa seringkali bingung untuk menentukan konsentrasi yang akan dipilih. Untuk itu diperlukan alat bantu yang akan memberikan dukungan dan kemudahan penentuan keputusan dalam memilih konsentrasi berdasarkan berbagai pertimbangan yang telah ditentukan.

Tujuan diadakannya penelitian ini adalah menghasilkan sebuah sistem pendukung keputusan memilih konsentrasi bagi mahasiswa. Melalui sistem pendukung keputusan ini diharapkan dapat membantu mengarahkan mahasiswa dalam memilih konsentrasi yang sesuai dengan minat dan kemampuannya. Penentuan konsentrasi berdasarkan jumlah sks yang telah diambil selama perkuliahan dan bobot nilai matakuliah terkait yang ada pada kurikulum yang ditawarkan. Studi kasus pada penelitian ini adalah penentuan konsentrasi pada Mahasiswa *International Class* Jurusan Teknik Informatika Angkatan 2013 STMIK AMIKOM Yogyakarta dengan menggunakan metode *K-Means Clustering Analysis*. Adapun Jurusan yang ditawarkan bagi mahasiswa jurusan Teknik Informatika adalah Pemrograman, Multimedia dan Jaringan Komputer.

Analisis Cluster merupakan suatu metode pengelompokan berdasarkan ukuran kedekatan (kemiripan) dan kemudian menggunakan karakteristik tersebut sebagai “vektor karakteristik” atau “centroid” [1]. Metode *K-Means Cluster Analysis* bisa menjadi

solusi untuk pengklasifikasian karakteristik dari objek. Algoritma *K-Means* memiliki ketelitian yang cukup tinggi terhadap ukuran objek, sehingga algoritma ini relatif lebih terukur dan efisien untuk pengolahan objek dalam jumlah besar. Selain itu algoritma *K-Means* ini tidak terpengaruh terhadap urutan objek.

Salah satu tahapan penting dalam menerapkan metode *K-means Cluster* adalah menentukan *centroid*, banyaknya *cluster* dan jarak *centroid*. Dengan membentuk beberapa *cluster* menggunakan algoritma *k-means* dapat juga mengetahui jarak antara *cluster* pusat (*centroid*) pada data yang akan dianalisa. Hasil ini menjadi dasar untuk mengklasifikasi data baru yang kemudian muncul sehingga diketahui kelompoknya[2].

Penentuan nilai awal centroid dapat dilakukan secara acak atau ditentukan dengan

$$d_{ij} = \sqrt{\sum_{k=1}^p \{x_{ik} - x_{jk}\}^2} \dots\dots\dots (1)$$

Dimana :

- Dij = Jarak objek antara objek i dan j
- P = Dimensi data
- Xik = Koordinat dari obyek i pada dimensi k
- Xjk = Koordinat dari obyek j pada dimensi k

Tahapan metode pada penelitian ini adalah :

- (1) Metode Pengumpulan Data
- (2) Metode Analisis
- (3) Pengolahan Data

2. Pembahasan

2.1 Metode Pengumpulan Data

Dalam membuat Sistem Pendukung Keputusan Pemilihan Konsentrasi, hal yang pertama dilakukan adalah pengumpulan data terkait Indeks Prestasi semester 3 mahasiswa *International Class* Jurusan Teknik Informatika Angkatan 2013 STMIK AMIKOM Yogyakarta. Selain indeks Prestasi, jumlah sks yang telah diambil juga menjadi salah satu syarat untuk mengambil konsentrasi. Namun dalam hal ini hanya menggunakan Indeks Prestasi mahasiswa sebagai data yang akan diolah. Adapun data yang telah diambil untuk digunakan pada penelitian ini dapat dilihat pada tabel berikut.

Tabel 1. Tabel Indeks Prestasi Mahasiswa Semester 3

Nama Mhs	Pemrograman		Multimedia		Jaringan Komputer	
	SBD	PWL	MM	Komgraf	HS II	Jarkom
Abdan	A	A	A	A	A	A
Falah	A	A	A	A	B	A

Wirda	A	A	A	A	A	B
Zaid	A	A	A	A	A	B
Fatih	B	A	A	A	A	A
Nanda	B	A	A	A	A	A
Surya	A	A	A	A	A	A
Josi	A	A	A	A	A	A
Elsa	B	A	A	A	A	B
Zainal	C	B	B	B	B	A
Dewi	B	A	A	A	A	B
Tyar	A	A	A	A	A	A

Keterangan :

- SBD = Sistem Basis Data
- PWL = Pemrograman Web Lanjut
- MM = Multimedia
- Komgraf = Komputer Grafis
- HS II = Hardware/Software II
- Jarkom = Jaringan Komputer

Dimana matakuliah Sistem Basis Data dan Pemrograman Web Lanjut merupakan matakuliah terkait konsentrasi Pemrograman, matakuliah Multimedia dan Komputer Grafis merupakan matakuliah terkait konsentrasi Multimedia kemudian matakuliah Hardware/Software II dan Jaringan Komputer merupakan matakuliah terkait konsentrasi Jaringan Komputer.

2.2. Metode Analisis

Pada tahap ini, agar data dapat diolah menggunakan algoritma *K-Means Clustering*, maka dilakukan proses perubahan data. Variabel yang diubah adalah nilai indeks prestasi yang dikelompokkan menjadi 3 kelompok, pertama untuk A ditransformasikan dengan nilai 3, B ditransformasikan dengan nilai 2 dan C ditransformasikan dengan nilai 1. Transformasi data dapat dilihat pada tabel berikut.

Tabel 2. Tabel Transformasi Indeks Prestasi

Nama Mhs	Pemrograman		Multimedia		Jaringan Komputer	
	SBD	PWL	MM	Komgraf	HS II	Jarkom
Abdan	3	3	3	3	3	3
Falah	3	3	3	3	2	3
Wirda	3	3	3	3	3	2
Zaid	3	3	3	3	3	2
Fatih	2	3	3	3	3	3
Nanda	2	3	3	3	3	3
Surya	3	3	3	3	3	3
Josi	3	3	3	3	3	3
Elsa	2	3	3	3	3	2
Zainal	1	2	2	2	2	3
Dewi	2	3	3	3	3	2
Tyar	3	3	3	3	3	3

2.3 Pengolahan Data

Pada tahap pengolahan data, data diolah untuk mendapatkan hasil yang kemudian dapat berguna dalam pengambilan keputusan. Secara umum tahapan algoritma k-means adalah :

1. Menentukan banyaknya cluster (k).
2. Menentukan centroid secara acak
3. Menghitung jarak centroid
4. Setelah dilakukan iterasi apakah centroidnya berubah?
 - Jika ya, hitung jarak data dari centroid
 - Jika tidak, selesai. Iterasi dihentikan.
5. Mengelompokkan data berdasarkan jarak yang terdekat.

Sehingga hal yang harus dilakukan pertama adalah menentukan banyaknya cluster yang diinginkan, dalam hal ini cluster (k) = 3. Kemudian penentuan centroid, dalam hal ini penentuan centroid dilakukan secara acak. Tabel berikut menunjukkan centroid.

Tabel 3. Tabel Centroid

Centroid	Nama Mhs	S B D	P W L	M M	K o m g r a f	H S I I	Jarkom
C1	Abdan	A	A	A	A	A	A
C2	Dewi	B	A	A	A	A	B
C3	Zainal	C	B	B	B	B	A

Agar dapat diolah, maka dilakukan transformasi sesuai dengan ketentuan yang telah ditetapkan pada tahap Metode Analisis.

Tabel 4. Tabel Centroid

Centroid	Nama Mhs	S B D	P W L	M M	K o m g r a f	H S I I	Jarkom
C1	Abdan	3	3	3	3	3	3
C2	Dewi	2	3	3	3	3	2
C3	Zainal	1	2	2	2	2	3

Kemudian menghitung jarak antara centroid. Dalam penelitian ini digunakan metode *hard k-means* untuk mengalokasikan setiap data ke dalam suatu *cluster* yang memiliki jarak paling dekat dengan titik pusat setiap *cluster*. Untuk mengetahui *cluster* mana yang paling

dekat dengan data, maka perlu dihitung jarak setiap data dengan titik pusat setiap *cluster*. Proses iterasi dapat dilihat pada tabel berikut.

Tabel 5. Tabel Hasil Menghitung Jarak, Iterasi 1

Nama Mhs	S B D	P W L	M M	K o m g r a f	H S I I	J a r k o m	Jarak Ke			Jarak ke cluster
							c 1	c 2	c 3	
Abdan	3	3	3	3	3	3	0	1	2	C1
Falah	3	3	3	3	2	3	1	1	2	C1
Wirda	3	3	3	3	3	2	1	1	3	C1
Zaid	3	3	3	3	3	2	1	1	3	C1
Fatih	2	3	3	3	3	3	1	1	2	C1
Nanda	2	3	3	3	3	3	1	1	2	C1
Surya	3	3	3	3	3	3	0	1	2	C1
Josi	3	3	3	3	3	3	0	1	2	C1
Elsa	2	3	3	3	3	2	1	0	2	C2
Zainal	1	2	2	2	2	3	2	2	0	C3
Dewi	2	3	3	3	3	2	1	0	2	C2
Tyar	3	3	3	3	3	3	0	1	2	C1

Setelah dilakukan iterasi 1, dilakukan pengelompokan data. Berikut merupakan tabel pengelompokan hasil iterasi 1.

Tabel 6. Tabel Pengelompokan Hasil Iterasi 1

Nama Mhs	C1	C2	C3
Abdan	1	0	0
Falah	1	0	0
Wirda	1	1	0
Zaid	1	1	0
Fatih	1	0	0
Nanda	1	0	0
Surya	1	0	0
Josi	1	0	0
Elsa	0	1	0
Zainal	0	0	1
Dewi	0	1	0
Tyar	1	0	0

Tahapan selanjutnya adalah menentukan centroid baru. Centroid baru tersebut digunakan untuk melihat apakah terdapat perubahan pada hasil iterasi 2. Jika tidak terdapat perubahan, maka iterasi dihentikan. Namun apabila terdapat perubahan, maka akan dilakukan iterasi berikutnya hingga tidak terdapat perubahan pada hasil iterasi tersebut[3]. Berikut merupakan centroid yang ditentukan secara acak.

Tabel 7. Tabel Centroid untuk Iterasi 2

Centroid	Nama Mhs	SBD	PWL	MM	Komgraf	HS II	Jarkom
C1	Surya	A	A	A	A	A	A
C2	Elsa	B	A	A	A	A	B
C3	Zaid	A	A	A	A	A	B

Sama halnya dengan centroid pada iterasi 1, agar dapat diolah maka data harus ditransformasikan sesuai dengan ketentuan pada tahap Metode Analisis.

Tabel 8. Tabel Centroid untuk Iterasi 2

Centroid	Nama Mhs	SBD	PWL	MM	Komgraf	HS II	Jarkom
C1	Surya	3	3	3	3	3	3
C2	Elsa	2	3	3	3	3	2
C3	Zaid	3	3	3	3	3	2

Setelah ditentukan centroid untuk iteasi kedua, maka ulangi langkah ketiga yaitu menentukan jarak centroid. Berikut merupakan hasil iterasi ke-2.

Iterasi 2

Nama Mhs	Pemrograman		Multimedia		Jaringan Komputer		Jarak Ke			Jarak terdekat ke cluster
	SBD	PWL	MM	Komgraf	HS II	Jarkom	c1	c2	c3	
Abdan	3	3	3	3	3	3	0	1	1	1
Falah	3	3	3	3	2	3	1	1,414214	1,414214	1
Wirda	3	3	3	3	3	2	1	0	1,414214	2
Zaid	3	3	3	3	3	2	1	0	1,414214	2
Fatih	2	3	3	3	3	3	1	1,414214	0	3
Nanda	2	3	3	3	3	3	1	1,414214	0	3
Surya	3	3	3	3	3	3	0	1	1	1
Josi	3	3	3	3	3	3	0	1	1	1
Elsa	2	3	3	3	3	2	1,414214	1	1	2
Zainal	1	2	2	2	2	3	2,828427	3	2,236068	3
Dewi	2	3	3	3	3	2	1,414214	1	1	2
Tyar	3	3	3	3	3	3	0	1	1	1

Centroid Iterasi 2

Nama Mhs	SBD	PWL	MM	Komgraf	HS II	Jarkom
Tyar	3	3	3	3	3	3
Wirda	3	3	3	3	3	2
Nanda	2	3	3	3	3	3

Gambar 1. Iterasi 2

Setelah dilakukan iterasi 2, dilakukan pengelompokan data. Berikut merupakan tabel pengelompokan hasil iterasi 1.

Tabel 9. Tabel Pengelompokan Hasi Iterasi 2

Nama Mhs	C1	C2	C3
Abdan	1	0	0
Falah	1	0	0
Wirda	0	1	0
Zaid	0	1	0
Fatih	0	0	1
Nanda	0	0	1
Surya	1	0	0
Josi	1	0	0
Elsa	0	1	1
Zainal	0	0	1
Dewi	0	1	1
Tyar	1	0	0

Setelah dilakukan iterasi 2 dapat dilihat bahwa terjadi perubahan data pada tabel pengelompokan hasil iterasi. Gambar berikut menunjukkan perubahan setelah iterasi 2.

Nama Mhs	Iterasi 1			Iterasi 2		
	C1	C2	C3	C1	C2	C3
Abdan	1	0	0	1	0	0
Falah	1	0	0	1	0	0
Wirda	1	1	0	0	1	0
Zaid	1	1	0	0	1	0
Fatih	1	0	0	0	0	1
Nanda	1	0	0	0	0	1
Surya	1	0	0	1	0	0
Josi	1	0	0	1	0	0
Elsa	0	1	0	0	1	1
Zainal	0	0	1	0	0	1
Dewi	0	1	0	0	1	1
Tyar	1	0	0	1	0	0

Gambar 2. Perubahan Setelah Iterasi 2

Karena terdapat perubahan, maka dilakukan iterasi berikutnya hingga tidak terdapat perubahan pada pengelompokan hasil iterasi. Maka langkah selanjutnya adalah kembali menentukan centroid dan menghitung jarak centroid, begitu seterusnya.

Centroid Iterasi 3

Nama Mhs	SBD	PWL	MM	Komgraf	HS II	Jarkom
Josi	3	3	3	3	3	3
Zaid	3	3	3	3	3	2
Fatih	2	3	3	3	3	3

Gambar 3. Centroid Iterasi 3

Iterasi 3

Nama Mhs	Pemrograman		Multimedia		Jaringan Komputer		Jarak Ke			Jarak terdekat ke cluster
	SBD	PWL	MM	Komgraf	HS II	Jarkom	c1	c2	c3	
Abdan	3	3	3	3	3	3	0	1	1	1
Falah	3	3	3	3	2	3	1	1,414214	1,414214	1
Wirda	3	3	3	3	3	2	1	0	1,414214	2
Zaid	3	3	3	3	3	2	1	0	1,414214	2
Fatih	2	3	3	3	3	3	1	1,414214	0	3
Nanda	2	3	3	3	3	3	1	1,414214	0	3
Surya	3	3	3	3	3	3	0	1	1	1
Josi	3	3	3	3	3	3	0	1	1	1
Elsa	2	3	3	3	3	2	1,414214	1	1	2
Zainal	1	2	2	2	2	3	2,828427	3	2,236068	3
Dewi	2	3	3	3	3	2	1,414214	1	1	2
Tyar	3	3	3	3	3	3	0	1	1	1

Gambar 4. Iterasi 3

Nama Mhs	Iterasi 1			Iterasi 2			Iterasi 3		
	C1	C2	C3	C1	C2	C3	C1	C2	C3
Abdan	1	0	0	1	0	0	1	0	0
Falah	1	0	0	1	0	0	1	0	0
Wirda	1	1	0	0	1	0	0	1	0
Zaid	1	1	0	0	1	0	0	1	0
Fatih	1	0	0	0	0	1	0	0	1
Nanda	1	0	0	0	0	1	0	0	1
Surya	1	0	0	1	0	0	1	0	0
Josi	1	0	0	1	0	0	1	0	0
Elsa	0	1	0	0	1	1	0	1	1
Zainal	0	0	1	0	0	1	0	0	1
Dewi	0	1	0	0	1	1	0	1	1
Tyar	1	0	0	0	0	0	0	0	0

Gambar 5. Perubahan Setelah Iterasi 3

Karena setelah iterasi 3 tidak terdapat perubahan, maka iterasi dihentikan. Dan dapat ditemukan hasilnya bahwa 4 dari 12 mahasiswa diarahkan untuk mengambil konsentrasi Pemrograman dan 4 mahasiswa diarahkan untuk mengambil konsentrasi Multimedia sementara 3-5 mahasiswa diarahkan untuk mengambil konsentrasi Jaringan Komputer. Apabila terdapat 2 pilihan seperti yang terjadi pada Elsa dan Dewi, maka dianjurkan untuk mengambil pilihan pertama yaitu C2.

3. Kesimpulan

Dari hasil penelitian yang dilakukan, maka penulis dapat menarik beberapa kesimpulan, antara lain sebagai berikut :

1. Pengujian yang dilakukan dalam penelitian ini, iterasi *clustering* data mahasiswa terjadi sebanyak 3 kali iterasi.
2. Berdasarkan dari hasil *Cluster* jarak terdekat jika menunjukkan 1 maka diarahkan untuk konsentrasi Programming, 2 Multimedia dan 3 Jaringan Komputer
3. Setelah dilakukan pengolahan data dan mengalami 3 kali iterasi *clustering* maka ditemukan hasil sebagai berikut, 4 dari 12 mahasiswa diarahkan untuk mengambil konsentrasi Pemrograman dan 4 mahasiswa diarahkan untuk mengambil konsentrasi Multimedia sementara 3 -5 mahasiswa diarahkan untuk mengambil konsentrasi Jaringan Komputer. Apabila terdapat 2 pilihan seperti yang terjadi pada Elsa dan Dewi, maka

dianjurkan untuk mengambil pilihan pertama yaitu C2, dimana C2 adalah konsentrasi Multimedia.

4. Hasil *cluster* juga dipengaruhi dari nilai *centroid* awal yang dipakai dan jumlah data yang dipakai, perbedaan pengambilan data pusat *centroid* awal yang dipakai juga akan mempengaruhi hasil *centroid* akhirnya.

Daftar Pustaka

- [1] Tanjung Muhammad Rusdi, Erlinda Ningsih, "Perbandingan Penggunaan Data Real dan Data Hasil Normalisasi Pada Pengelompokan Data Dengan Menggunakan Metode Clustering (Studi Kasus PT. PHP Motor)", *Prosiding KeTIK 2014* UIN. Universitas Islam Negeri Medan – Sumatera Utara
- [2] Ediyanto, Muhlasah Novitasari Mara, Neva Satyahadewi, Pengklasifikasian Karakteristik Dengan Metode K-Means Cluster Analysis. *Buletin Ilmiah Mat. Stat. dan Terapannya (Bimaster)*, Volume 02, no 2, 133-136. 2013
- [3] Agusta, Y. K-Means-Penerapan, Permasalahan dan Metode Terkait *Jurnal Sistem dan Informatika*, Vol. 3, 47-60. Februari 2007

Biodata Penulis

Josi Aranda, Mahasiswa Jurusan Teknik Informatika (BCIT) Angkatan 2013. Saat ini merupakan mahasiswa semester V 13-BCIT-01 STMIK AMIKOM Yogyakarta.

Wirda Astari Galvani Natasya, Mahasiswi Jurusan Teknik Informatika (BCIT) Angkatan 2013. Saat ini merupakan mahasiswi semester V 13-BCIT-01 STMIK AMIKOM Yogyakarta.

