

Pengelompokan Peminatan Outline Tugas Akhir Dengan Menggunakan Algoritma K-Means Pada AMIK MI BSI Jakarta

Achmad Maezar Bayu Aji

Sekolah Tinggi Manajemen dan Ilmu Komputer Nusa Mandiri Jakarta
achmad.azb@bsi.ac.id

ABSTRAK - Sebelum membuat tugas akhir, mahasiswa memilih outline untuk menentukan minat dan tujuan dari pembuatan tugas akhir tersebut. Pada AMIK MI BSI Jakarta mempunyai banyak cabang dan banyak peminatan outline. Salah satu masalah dalam pemilihan outline tugas akhir di AMIK MI BSI Jakarta adalah terlalu banyaknya data peminatan outline dan cabang. Tujuan dari penelitian ini untuk mengetahui dan mengelompokkan outline pendaftaran tugas akhir yang paling banyak diminati dan yang paling sedikit diminati untuk nantinya akan diperbaharui lagi peminatan outline tugas akhir di AMIK MI BSI Jakarta. Algoritma yang digunakan untuk mengelompokkan data tersebut menggunakan algoritma k-means K-Means adalah suatu metode penganalisaan data. Metode k-means berusaha mengelompokkan data yang ada ke dalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama satu sama lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada di dalam kelompok yang lain. Atribut yang digunakan dalam pengelompokan peminatan outline tugas akhir adalah jenis outline pendaftaran tugas akhir, cabang AMIK MI BSI Jakarta, nilai tugas akhir mahasiswa. Cluster yang terbentuk ada 3 cluster, cluster pertama mempunyai 9 anggota data mahasiswa, cluster kedua mempunyai 8 anggota data mahasiswa, cluster ketiga mempunyai anggota data mahasiswa

Kata Kunci: *Tugas Akhir, K-Means, Peminatan Outline*

ABSTRAK - Sebelum membuat tugas akhir, mahasiswa memilih outline untuk menentukan minat dan tujuan dari pembuatan tugas akhir tersebut. Pada AMIK MI BSI Jakarta mempunyai banyak cabang dan banyak peminatan outline. Salah satu masalah dalam pemilihan outline tugas akhir di AMIK MI BSI Jakarta adalah terlalu banyaknya data peminatan outline dan cabang. Tujuan dari penelitian ini untuk mengetahui dan mengelompokkan outline pendaftaran tugas akhir yang paling banyak diminati dan yang paling sedikit diminati untuk nantinya akan diperbaharui lagi peminatan outline tugas akhir di AMIK MI BSI Jakarta. Algoritma yang digunakan untuk mengelompokkan data tersebut menggunakan algoritma k-means K-Means adalah suatu metode penganalisaan data. Metode k-means berusaha mengelompokkan data yang ada ke dalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama satu sama lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada di dalam kelompok yang lain. Atribut yang digunakan dalam pengelompokan peminatan outline tugas akhir adalah jenis outline pendaftaran tugas akhir, cabang AMIK MI BSI Jakarta, nilai tugas akhir mahasiswa. Cluster yang terbentuk ada 3 cluster, cluster pertama mempunyai 9 anggota data mahasiswa, cluster kedua mempunyai 8 anggota data mahasiswa, cluster ketiga mempunyai anggota data mahasiswa

Kata Kunci: *Tugas Akhir, K-Means, Peminatan Outline*

1. a Latar Belakang

Dalam penentuan kelulusan suatu universitas atau akademik salah satu faktor yang menentukan adalah tugas akhir. Tugas akhir merupakan hasil tertulis dari pelaksanaan suatu penelitian, yang dibuat untuk pemecahan masalah tertentu dengan menggunakan kaidah-kaidah yang berlaku dalam bidang ilmu tersebut.

Syarat untuk pendaftaran tugas akhir di AMIK MI BSI Jakarta, mahasiswa tersebut harus menyelesaikan 110 sks dan tidak ada mata kuliah unggulan yang di her, dan ipk minimal 2,75. Kemudian, mahasiswa memilih *outline* untuk menentukan minat dan tujuan dari pembuatan tugas akhir tersebut. Di AMIK MI BSI Jakarta untuk *outline* tugas akhir diantaranya adalah : Jaringan Komputer,

Pemrograman - Berorientasi Bisnis - Web Programming, Pemrograman - Berorientasi Science, Pemrograman - Animasi Interaktif, Pembuatan Alat, Pemrograman Berorientasi Bisnis Visual, Perancangan Sistem Berorientasi Objek, Pemrograman Berorientasi Objek, Aplikasi Program Akuntansi, Perancangan Sistem, Proyek Sistem Informasi Berorientasi Objek.

Dan untuk AMIK MI BSI Jakarta dibagi dalam beberapa cabang, yaitu cabang margonda, fatmawati, Kalimalang, Kramat-Salemba, Jatiwaringin, Ciledug, Wolter Monginsidi, Cengkareng, Pondok Aren, Dewi Sartika, Pemuda (Rawamangun).

1.b Rumusan Masalah

Salah satu masalah dalam pemilihan *outline* tugas akhir di AMIK MI BSI Jakarta adalah terlalu banyaknya data peminatan *outline* dan cabang. Sehingga dalam penelitian ini fokus pada metode data mining dengan kasus pengelompokan data (*clustering*). Menurut Chandra, dkk (2014) “Adanya data dalam skala besar memungkinkan metode data mining dengan teknik clustering yang dapat mengelompokkan data ke dalam beberapa kelompok yang diinginkan. Teknik clustering yang digunakan yaitu K-Means”[1].

1.c Batasan Masalah

Penelitian ini akan melakukan pengelompokan data berdasarkan pada data mahasiswa AMIK MI BSI Jakarta dengan data bimbingan periode 2014-1 di seluruh cabang AMIK MI BSI Jakarta. Atribut yang digunakan dalam pengelompokan peminatan *outline* tugas akhir adalah jenis *outline* pendaftaran tugas akhir, cabang AMIK MI BSI Jakarta, nilai tugas akhir mahasiswa.

1.d Tujuan

Tujuan dari penelitian ini untuk mengetahui dan mengelompokan *outline* pendaftaran tugas akhir yang paling banyak diminati dan yang paling sedikit diminati untuk nantinya akan diperbaharui lagi peminatan *outline* tugas akhir di AMIK MI BSI Jakarta.

1.e Manfaat Penelitian

Manfaat dari penelitian ini mahasiswa bisa lebih fokus dalam memilih *outline* tugas akhir karena sudah mengetahui *outline* yang paling banyak diminati dan mana yang sedikit, juga untuk institusi lebih diperbaharui dan direvisi kembali *outline* pendaftaran tugas akhir mahasiswa

1.f Metode Penelitian

Metode yang digunakan dalam penelitian ini dengan menggunakan algoritma k-means.

Menurut Ong (2013) “Dengan menggunakan metode ini, data data yang telah didapatkan dapat dikelompokkan kedalam beberapa cluster berdasarkan kemiripan dari data-data tersebut, sehingga data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu cluster dan yang memiliki karakteristik yang berbeda dikelompokkan dalam cluster yang lain yang memiliki karakteristik yang sama”. [7].

Menurut Budi (2007) langkah-langkah melakukan clustering dengan metode K-Means adalah sebagai berikut[3]:

- a. Pilih jumlah cluster k.
- b. Inialisasi k pusat cluster ini bisa dilakukan dengan berbagai cara. Namun yang paling sering dilakukan adalah dengan cara random. Pusat-pusat cluster diberi nilai awal dengan angka-angka random,
- c. Alokasikan semua data/ objek ke cluster terdekat. Kedekatan dua objek ditentukan berdasarkan jarak kedua objek tersebut. Demikian juga kedekatan suatu data ke cluster tertentu ditentukan jarak antara data dengan pusat cluster. Dalam tahap ini perlu dihitung jarak tiap data ke tiap pusat cluster. Jarak paling antara satu data dengan satu cluster tertentu akan menentukan suatu data masuk dalam cluster mana. Untuk menghitung jarak semua data ke setiap titik pusat cluster dapat menggunakan teori jarak Euclidean yang dirumuskan sebagai berikut:

$$D(i,j)=\text{Error! Reference source not}$$

found.Error! Reference source not found.Error!

Reference source not found.

Dimana:

$D_{(i,j)}$ = Jarak data ke i ke pusat cluster j

X_{ki} = Data ke i pada atribut data ke k

X_{kj} = Titik pusat ke j pada atribut ke k **Error!**

Reference source not found.

- d. Hitung kembali pusat cluster dengan keanggotaan cluster yang sekarang. Pusat cluster adalah rata-rata dari semua data/ objek dalam cluster tertentu. Jika dikehendaki bisa juga menggunakan median dari cluster tersebut. Jadi rata-rata (mean) bukan satu-satunya ukuran yang bisa dipakai.

Tugaskan lagi setiap objek memakai pusat cluster yang baru. Jika pusat cluster tidak berubah lagi maka proses clustering selesai. Atau, kembali ke langkah nomor 3 sampai pusat cluster tidak berubah lagi.

2.a Dasar Teori Data Mining

Menurut Kusriani dan Lutfi (2009)

“Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terakut dari berbagai database besar”[2].

Dimana hasil dari proses penggalian tersebut akan membentuk pola-pola dari kumpulan data, yang sering disebut dengan pengenalan pola (pattern recognition). Pengenalan pola merupakan bagian dari data mining. Menurut Budi (2007) "Data mining sering juga disebut Knowledge Discovery in Database (KDD), yaitu kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar" [3].

Clustering

Menurut Zaki (2014) "Clustering adalah suatu tugas yang membagi poin menjadi beberapa kelompok-kelompok alami yang disebut klaster, sehingga poin-poin dalam kelompok sangat mirip, sedangkan poin-poin yang diluar cluster sedikit berbeda. Tergantung pada data dan karakteristik klaster yang diinginkan, ada berbagai jenis pengelompokan paradigma seperti berbasis perwakilan, hirarkis, berdasarkan kepadatan, berdasarkan grafik dan spektral clustering" [4].

Data mining adalah suatu metode pengolahan data untuk menemukan pola yang tersembunyi dari data tersebut. Hasil dari pengolahan data dengan metode data mining ini dapat digunakan untuk mengambil keputusan di masa depan. Menurut Budi(2007) "Data mining ini juga dikenal dengan istilah pattern recognition dan Data mining merupakan metode pengolahan data berskala besar oleh karena itu data mining ini memiliki peranan penting dalam bidang industri, keuangan, cuaca, ilmu dan teknologi. Secara umum kajian data mining membahas metode-metode seperti, clustering, klasifikasi, regresi, seleksi variable, dan market basket analisis" [3].

K-Means

Menurut Kusumadewi dan Purnomo (2010) "K-Means adalah suatu teknik pengelompokan data yang mana keberadaan tiap-tiap titik data dalam suatu cluster ditentukan oleh derajat keanggotaan. Teknik ini pertama

kali diperkenalkan oleh Jim Bezdek pada tahun 1981" [5].

K-Means adalah suatu metode penganalisaan data. Metode k-means berusaha mengelompokkan data yang ada ke dalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama satu sama lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada di dalam kelompok yang lain. Dengan kata lain, metode ini berusaha untuk meminimalkan variasi antar data yang ada di dalam suatu cluster dan memaksimalkan variasi dengan data yang ada di cluster lainnya.

Menurut Sri (2007) William membagi algoritma clustering ke dalam kelompok besar seperti berikut [6] :

1. Partitioning algorithms: algoritma dalam kelompok ini membentuk bermacam partisi dan kemudian mengevaluasinya dengan berdasarkan beberapa kriteria.
2. Hierarchy algorithms: pembentukan dekomposisi hirarki dari sekumpulan data menggunakan beberapa kriteria.
3. Density- based: pembentukan cluster berdasarkan pada koneksi dan fungsi densitas.
4. Grid- based: pembentukan cluster berdasarkan pada struktur multiple level granularity
5. Model- based: sebuah model dianggap sebagai hipotesa untuk masing- masing cluster dan model yang baik dipilih diantara model hipotesa tersebut.

3. HASIL DAN PEMBAHASAN

Dalam penelitian diperoleh data dari bagian akademik AMIK MI BSI Jakarta berupa data bimbingan periode 2014-1 dan data hasil lulus sidang. Data-data tersebut berisi tentang data mahasiswa bimbingan, dan dalam penelitian ini atribut yang digunakan hanya sebatas, nim, nama mahasiswa, cabang, outline tugas akhir dan nilai. Berikut ini contoh adalah data mahasiswa bimbingan periode 2014-1 yang ditampilkan dalam tabel 1.

Tabel 1. Contoh Data Mahasiswa Bimbingan Periode 2014-1 AMIK MI BSI Jakarta

No	Nim	Nama Mahasiswa	kd_cabang	Kd Outline	Nama Outline	NILAI TA
1	12112979	MIFTACHUL BASORI	1	98	Pemrograman Berorientasi Bisnis Visual (KA & MI)	B
2	12112251	RESTI SEPRIYANI	1	108	Perancangan Sistem (Dekstop/Web)	B
3	18113933	PUSNANDAR	1	101	Pemrograman Berorientasi Objek (Desktop/Web)	B

4	1211775 8	MOCHAMAD RIFKI	2	42	Pemrograman - Berorientasi Bisnis - Web Programming (Untuk mahasiswa : AMIK)	A
5	1211770 2	ROBERTUS EKO PRASETYO	2	42	Pemrograman - Berorientasi Bisnis - Web Programming (Untuk mahasiswa : AMIK)	A
6	1811294 3	YEANNE SITI YANTO	7	98	Pemrograman Berorientasi Bisnis Visual (KA & MI)	A
7	1211227 1	NUR HAMIDAH	7	42	Pemrograman - Berorientasi Bisnis - Web Programming (Untuk mahasiswa : AMIK)	C
8	1811271 8	WULAN DARI	7	98	Pemrograman Berorientasi Bisnis Visual (KA & MI)	B
9	1811390 8	YULI NURMALITA	11	100	Perancangan Sistem Berorientasi Objek (Desktop/Web)	A
10	1811250 5	VIRGINIA RIA UNIKA NURMAWATI	11	42	Pemrograman - Berorientasi Bisnis - Web Programming (Untuk mahasiswa : AMIK)	A
11	1811129 0	AHMAD SOFIYAN	11	42	Pemrograman - Berorientasi Bisnis - Web Programming (Untuk mahasiswa : AMIK)	A
12	1811078 1	YESSY DEVIANTARI	11	101	Pemrograman Berorientasi Objek (Desktop/Web)	A
13	1811238 2	ASEP ZAKARIA	12	100	Perancangan Sistem Berorientasi Objek (Desktop/Web)	A
14	1211783 9	IRFAN MAULANA	12	42	Pemrograman - Berorientasi Bisnis - Web Programming (Untuk mahasiswa : AMIK)	A
15	1211243 6	WAHYUNINGSIH	31	42	Pemrograman - Berorientasi Bisnis - Web Programming (Untuk mahasiswa : AMIK)	A
16	1211243 0	AFRILIA RAHMAN	31	42	Pemrograman - Berorientasi Bisnis - Web Programming (Untuk mahasiswa : AMIK)	A
17	1211006 7	ASEP SAEPUDIN	31	108	Perancangan Sistem (Dekstop/Web)	A
18	1811105 0	SOPIAH SITI NURAENI	31	98	Pemrograman Berorientasi Bisnis Visual (KA & MI)	B
19	1811368 4	NANANG RIYADI	32	98	Pemrograman Berorientasi Bisnis Visual (KA & MI)	A
20	1211781 7	AHMAD FAUZI	32	42	Pemrograman - Berorientasi Bisnis - Web Programming (Untuk mahasiswa : AMIK)	A

Untuk pengolahan data dalam algoritma k-means, harus dalam berbentuk angka, sehingga data yang menggunakan nominal diganti dengan angka. Dalam data tersebut, data yang bersifat nominal seperti nama mahasiswa, nama cabang, dan nama outline, serta nilai akan diganti. Untuk nama mahasiswa akan diganti dengan nim, nama cabang diganti dengan kode cabang, nama outline diganti dengan kode outline, dan untuk nilai ta akan dibuat inisialisasinya dengan tabel di bawah ini.

Tabel 2. Inisialisasi tabel nilai

NILAI TA	NILA I
A	1
B	2
C	3
D	4

E 5

Setelah semua data sudah berbentuk angka, kemudian data-data tersebut diolah menggunakan algoritma *k-means clustering*. Untuk mengelompokan data tersebut perlu dilakukan dalam beberapa tahap. Untuk tahapan-tahapan dalam algoritma k-means clustering adalah sebagai berikut:

1. Langkah pertama dalam menggunakan algoritma k-means adalah dengan menentukan jumlah cluster yang diinginkan. Dalam penelitian ini jumlah cluster yang digunakan untuk mengelompokan data berjumlah 3 cluster.
2. Memberikan nilai pada titik-titik pusat cluster atau k pusat cluster, untuk memberikan nilai pada k pusat cluster ini bisa dilakukan

dengan berbagai cara, dalam penelitian ini untuk menentukan pusat-pusat cluster diberi nilai awal dengan angka-angka random. Berikut ini adalah titik pusat-pusat cluster pada tabel 3.

Tabel 3. Titik Pusat Awal Cluster

No	Nim	kd_cabang	Kd Outline	NILAI TA
1	12117758	2	42	1
2	18110781	11	101	1
3	18111050	31	98	2

- Menempatkan data-data ke dalam setiap cluster. Untuk menempatkan data-data ke dalam setiap cluster perlu dihitung jarak tiap data ke tiap pusat cluster. Jarak paling antara satu data dengan satu cluster tertentu akan menentukan suatu data masuk dalam cluster mana. dalam penelitian ini Untuk menghitung jarak semua data ke setiap titik pusat cluster menggunakan teori jarak Euclidean. Berikut ini adalah contoh perhitungan jarak dari data mahasiswa pertama ke setiap titik pusat cluster:
 Untuk perhitungan mahasiswa pertama di cluster pertama.
 Titik mahasiswa pertama:(1,98,2)
 Untuk titik pusat cluster pertama : (2, 42, 1).
 Perhitungan :
 $D_{(1,1)} = \sqrt{(1-2)^2 + (98-42)^2 + (2-1)^2}$
 Hasil : 56,01
 Jadi, jarak dari data mahasiswa pertama dengan titik pusat cluster pertama adalah 56,01

Untuk perhitungan mahasiswa pertama di cluster kedua.
 Titik mahasiswa pertama:(1, 98, 2)
 Untuk titik pusat cluster kedua : (11, 101, 1).
 Perhitungan :
 $D_{(1,1)} = \sqrt{(1-11)^2 + (98-101)^2 + (2-1)^2}$
 Hasil : 10,48
 Jadi, jarak dari data mahasiswa pertama dengan titik pusat cluster kedua adalah 10,48

Untuk perhitungan mahasiswa pertama di cluster ketiga.
 Titik mahasiswa pertama:(1, 98, 2)
 Untuk titik pusat cluster ketiga : (31,98,2).
 Perhitungan :
 $D_{(1,1)} = \sqrt{(1-31)^2 + (98-98)^2 + (2-2)^2}$
 Hasil : 30
 Jadi, jarak dari data mahasiswa pertama dengan titik pusat cluster ketiga adalah 30

Berdasarkan hasil perhitungan dari ketiga titik cluster tersebut, dapat kita ketahui bahwa jarak data mahasiswa yang pertama dengan ketiga titik cluster tersebut yang paling dekat jaraknya adalah dengan titik pusat cluster yang kedua yaitu sebesar 10,48 jadi dapat kita simpulkan bahwa data mahasiswa pertama masuk kedalam cluster pertama. Berikut ini contoh untuk perhitungan data-data mahasiswa yang ditunjukkan pada tabel 4.

Tabel 4. Contoh Hasil Perhitungan Data Mahasiswa Ke Setiap Titik Cluster

No	Nim	kd_cabang	Kd Outline	NILAI TA	jarak ke cluster			jarak terdekat dengan cluster
					C1	C2	C3	
1	12112979	1	98	2	56,01785	10,48809	30	2
2	12112251	1	108	2	66,01515	12,24745	31,62278	2
3	18113933	1	101	2	59,01695	10,04988	30,14963	2
4	12117758	2	42	1	0	59,68249	63,07139	1
5	12117702	2	42	1	0	59,68249	63,07139	1
6	18112943	7	98	1	56,22277	5	24,02082	2
7	12112271	7	42	3	5,385165	59,16925	60,93439	1

8	18112718	7	98	2	56,23166	5,09902	24	2
9	18113908	11	100	1	58,69412	1	20,12461	2
10	18112505	11	42	1	9	59	59,47268	1
11	18111290	11	42	1	9	59	59,47268	1
12	18110781	11	101	1	59,68249	0	20,24846	2
13	18112382	12	100	1	58,85576	1,414214	19,13113	2
14	12117839	12	42	1	10	59,00847	59,14389	1
15	12112436	31	42	1	29	62,29767	56,00893	1
16	12112430	31	42	1	29	62,29767	56,00893	1
17	12110067	31	108	1	72,09022	21,18962	10,04988	3
18	18111050	31	98	2	63,07139	20,24846	0	3
19	18113684	32	98	1	63,52952	21,2132	1,414214	3
20	12117817	32	42	1	30	62,62587	56,01785	1

4. Menghitung kembali pusat cluster dengan keanggotaan cluster yang sudah terbentuk. titik pusat cluster yang baru adalah rata-rata dari semua anggota dalam setiap cluster tertentu. Berikut ini perhitungan untuk menentukan titik pusat cluster yang baru:

Untuk titik cluster pertama dari contoh data mahasiswa sebagai berikut:

Total anggota cluster pertama : 9 anggota

Total anggota cluster kedua : 8 anggota

Total anggota cluster ketiga : 3 anggota

Rata-rata cluster pertama

$$C1_{(1)} = (2+2+7+11+11+12+31+31+32)/9 = 15,44$$

$$C1_{(2)} = (42+42+42+42+42+42+42+42+42)/9 = 42$$

$$C1_{(3)} = (1+1+1+3+1+1+1+1+1)/9 = 1,22$$

Jadi, titik pusat baru untuk cluster pertama adalah (15.44, 42, 1.22)

Rata-rata cluster kedua

$$C2_{(1)} = (1+1+1+7+7+11+11+12)/8 = 6,38$$

$$C2_{(2)} = (98+108+101+98+98+100+101+100)/8 = 100,5$$

$$C2_{(3)} = (2+2+2+1+2+1+1+1+1)/8 = 1,75$$

Jadi, titik pusat baru untuk cluster pertama adalah (6.38, 100.5, 1.75)

Rata-rata cluster ketiga

$$C3_{(1)} = (31+31+32)/3 = 31,33$$

$$C3_{(2)} = (108+108+98)/3 = 104,67$$

$$C3_{(3)} = (1+2+1)/3 = 1,3$$

Jadi, titik pusat baru untuk cluster pertama adalah (31.33, 104.67, 1.3)

Tabel 5. Titik Pusat Cluster Baru

Clus ter	kd_cab ang	Kd Outline	NILAI TA
1	15,44	42	1,22
2	6,38	100,5	1,75
3	31,33	104,67	1,3

5. Hitung kembali setiap anggota memakai titik pusat cluster yang baru. Jika titik pusat cluster tidak berubah lagi maka proses clustering selesai. Atau, kembali ke langkah nomor 3 sampai pusat cluster tidak berubah lagi.

Berikut ini adalah contoh perhitungan data mahasiswa pertama dengan titik pusat cluster yang baru

Titik mahasiswa pertama: (1,98,2)

Untuk titik pusat baru cluster pertama : (15.44, 42, 1.22).

Perhitungan :

$$D_{(1,1)} = \sqrt{(1-15,44)^2 + (98-42)^2 + (2-1,22)^2}$$

Hasil : 57,83

Jadi, jarak dari data mahasiswa pertama dengan titik pusat cluster pertama adalah 57,83

Untuk perhitungan mahasiswa pertama di cluster kedua.

Titik mahasiswa pertama:(1,98,2)

Untuk titik baru pusat cluster kedua : (6.38, 100.5, 1.75).

Perhitungan :

$$D_{(1,2)} = \sqrt{(1-6,38)^2 + (98-100,5)^2 + (2-1,75)^2}$$

Hasil : 5,93772

Jadi, jarak dari data mahasiswa pertama dengan titik pusat cluster kedua adalah 5,93

Untuk perhitungan mahasiswa pertama di cluster ketiga.

Titik mahasiswa pertama:(1,98,2)

Untuk titik pusat baru cluster ketiga : (31.33, 104.67, 1.3).

Perhitungan :

$$D_{(1,3)} = \sqrt{(1-31,33)^2 + (98-104,67)^2 + (2-1,3)^2}$$

Hasil : 31,06

Jadi, jarak dari data mahasiswa pertama dengan titik pusat cluster ketiga adalah 31,06

Berdasarkan hasil perhitungan dari ketiga titik cluster tersebut, dapat kita ketahui bahwa jarak data mahasiswa yang pertama dengan ketiga titik cluster tersebut yang paling dekat jaraknya adalah dengan titik pusat cluster yang kedua yaitu sebesar 5,93 jadi dapat kita simpulkan bahwa data mahasiswa pertama masuk kedalam cluster kedua. Berikut ini contoh untuk perhitungan data-data mahasiswa yang ditunjukkan pada tabel 6.

Tabel 6. Contoh Hasil Perhitungan Data Mahasiswa Dengan Titik Pusat Cluster Baru

No	Nim	kd_cabang	Kd Outline	NILAI TA	jarak ke cluster			jarak terdekat dengan cluster
					C1	C2	C3	
1	12117758	2	42	1	13,4418	58,66853	69,19442	1
2	12117702	2	42	1	13,4418	58,66853	69,19442	1
3	12112271	7	42	3	8,625659	58,51664	67,24855	1
4	18112505	11	42	1	4,445447	58,68694	65,88572	1
5	18111290	11	42	1	4,445447	58,68694	65,88572	1
6	12117839	12	42	1	3,447028	58,77412	65,58405	1
7	12112436	31	42	1	15,56156	63,47406	62,67159	1
8	12112430	31	42	1	15,56156	63,47406	62,67159	1
9	12117817	32	42	1	16,56146	63,86859	62,6743	1
10	12112979	1	98	2	57,83703	5,937752	31,06264	2
11	12112251	1	108	2	67,56569	9,233466	30,52029	2
12	18113933	1	101	2	60,74637	5,408965	30,55925	2
13	18112943	7	98	1	56,63287	2,682704	25,2295	2
14	18112718	7	98	2	56,63781	2,587837	25,23743	2
15	18113908	11	100	1	58,17011	4,707112	20,86163	2
16	18110781	11	101	1	59,16724	4,707112	20,66078	2
17	18112382	12	100	1	58,10234	5,691827	19,88838	2

18	1211006 7	31	108	1	67,8097 5	25,7479 5	3,35973 2	3
19	1811105 0	31	98	2	58,1267 8	24,7478 7	6,71474 5	3
20	1811368 4	32	98	1	58,3976 2	25,7526 1	6,71027 6	3

Dalam penelitian ini, titik pusat cluster tidak berubah lagi dan data-data tidak ada lagi yang berpindah ke masing-masing cluster pada iterasi ke-3, iterasi clustering yang terjadi sebanyak 3 kali.

Hasil dari penelitian ini dengan mengelompokkan ke dalam 3 cluster dengan menggunakan algoritma k-means adalah sebagai berikut

1. Untuk cluster pertama berisi 9 anggota dengan data mahasiswanya terdiri dari 2 dari Fatmawati, Jatiwaringin dan Dewi Sartika. Untuk cabang Kramat-Salemba, Ciledug, dan Warung Jati hanya berisi 1 data mahasiswa. Outline peminatannya didominasi oleh Pemrograman - Berorientasi Bisnis - Web Programming dengan nilai tugas akhir dominan A dengan hanya satu mahasiswa yang mempunyai nilai C.
2. Untuk cluster kedua berisi 8 anggota dengan data mahasiswanya terdiri dari 3 mahasiswa cabang margonda, 2 mahasiswa cabang Kramat-Salemba dan Jatiwaringin, dan hanya 1 mahasiswa dari cabang Ciledug, dengan outline peminatannya terdiri dari Pemrograman Berorientasi Bisnis Visual, Perancangan Sistem Berorientasi Objek, Pemrograman Berorientasi Objek dan Perancangan Sistem dengan nilai tugas akhir dominan A dan B.
3. Untuk cluster ketiga hanya berisi 3 anggota dengan data mahasiswanya 2 dari cabang Dewi Sartika dan dari cabang Warung Jati, dengan outline peminatannya terdiri dari Pemrograman Berorientasi Bisnis Visual, Perancangan Sistem dengan nilai tugas akhir dominan A dan B.

4. PENUTUP

Kesimpulan

Kesimpulan yang didapat dari penelitian ini adalah sebagai berikut

1. Metode Algoritma K-Means digunakan untuk metode pengelompokan data dengan menggunakan titik pusat K, dimana untuk menentukan nilai K pada saat memproses data pertama kali umumnya dengan random.
2. Dalam menentukan titik pusat cluster K baru adalah dengan menghitung nilai rata-rata dalam setiap anggotanya.
3. Pengelompokan dalam penelitian memakai 3 cluster. Untuk cluster pertama berisi 9 data mahasiswa, cluster kedua berisi 8 data mahasiswa, dan untuk cluster ketiga berisi 3 data mahasiswa. Iterasi dalam penelitian ini hanya mencapai 3 iterasi.
4. Data yang digunakan dalam penelitian ini adalah data mahasiswa bimbingan periode 2014-1 dan untuk pengolahan datanya hanya menggunakan 20 data mahasiswa.

Saran

Dalam penelitian ini penulis memberikan saran-saran sebagai berikut

1. Outline pendaftaran tugas akhir yang paling dominan Pemrograman - Berorientasi Bisnis - Web Programming. Perlu ditingkatkan lagi untuk outline pendaftaran tugas akhir yang lain dengan melakukan sosialisasi atau membuat definisi yang jelas pada setiap masing-masing outline pendaftaran tugas akhir.
2. Menambah lagi metode yang digunakan, tidak hanya sebatas algoritma k-means. Untuk mengoptimasi nilai pusat K cluster yang umumnya random.
3. Atribut yang digunakan dalam penelitian seharusnya bisa ditambah dengan IPK, untuk penelitian selanjutnya.
4. Untuk pengolahan data selanjutnya data yang digunakan jangan hanya sebatas satu periode tapi seharusnya bisa lebih dari satu periode.

5. PUSTAKA

- [1] Candra Nugraha, Deka Dwinavinta., dkk. 2014. Klasterisasi Judul Buku

- dengan Menggunakan Metode K-Means. Jurnal SNATI, Juni 2014
- [2] Kusrini dan Lutfi, E.T. 2009. Algoritma Data Mining. Yogyakarta: Andi Offset.
- [3] Santosa, Budi. 2007. Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis. Yogyakarta: Graha Ilmu.
- [4] Zaki, Mohammed J., and Wagner Meira Jr. *Data mining and analysis: fundamental concepts and algorithms*. Cambridge University Press, 2014.
- [5] Kusumadewi dan Purnomo. 2010. Aplikasi Logika Fuzzy untuk mendukung keputusan. Yogyakarta: Graha Ilmu.
- [6] Andayani, Sri., 2007, Pembentukan Cluster dalam Knowledge Discovery in database dengan Algoritma K-Means, Semnas Matematika dan Pendidikan Matematika.
- [7] Ong, Johan Oscar, 2013. Implementasi Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing President University. JITI Vol.12, No.1.