

PENERAPAN METODE K-MEANS CLUSTERING UNTUK MENENTUKAN CALON MAHASISWA BERPRESTASI

Lidya Rizki Ananda,S.kom, M.kom
Magister Ilmu Komputer, Universitas Putra Indonesia “YPTK” Padang
Jl.Lubuk Begalung, Padang , Indonesia
*E-Mail : lidyarizkiananda@gmail.com

Dalam menentukan calon mahasiswa berprestasi tersebut membutuhkan waktu yang cukup lama untuk mendapatkan hasil yang baik, Seiring dengan perkembangan *software* pendukungnya yaitu dalam pengolahan data dan penyajian informasi, Untuk memenuhi kebutuhan-kebutuhan tersebut, salah satunya cara yaitu memanfaatkan dengan menerapkan *Data Mining*. Penggunaan teknik *Data Mining* diharapkan dapat membantu mempercepat proses pengambilan keputusan, Yang mana pada perguruan tinggi ini mengolah data nilai mahasiswa masih menggunakan secara manual dalam menentukan calon mahasiswa berprestasi tersebut. Dengan belum ada nya sistem pendukung pengambilan keputusan yang tepat, untuk mendapatkan informasi yang akurat menentukan calon mahasiswa berprestasi membutuhkan waktu yang cukup lama untuk mendapatkan hasil yang baik. Dengan menggunakan bidang Data Mining di iringi dengan Metode Clustering dengan menggunakan Algoritma K-Means dapat mempermudah pengambilan keputusan untuk melakukan pengclusteran data pada perguruan tinggi untuk mengelola informasi yang terkandung di dalam menentukan calon mahasiswa berprestasi menjadi sebuah pengetahuan (*knowledge*) yang baru dan dapat lebih kompetitif. Seperti pada *Akademi Manajemen Gunung Leuser Palas Sumatera Utara*.

Kata Kunci : Data Mining, Algoritma K-means, Clustering..

It takes such a long time to choose a prospective student achiever. Since the support software has been developed in data processing and presentation of information. On of the ways to solve the problem is by using data mining. Aplicating data mining aims to speed up the process of decision making, which university used to process the student data manually. Data mining is combined with clustering method by using K-Means algorithm can make the process easier to choose a prospective student achiever, then become a new knowledge and more competitive like for *Akademi Manajemen Gunung Leuser Palas Sumatera Sumatera Utara*.

Keywords : Data Mining, K-means Algorithm, Clustering

1. Pendahuluan

Perkembangan komputer pada saat ini semakin meningkat menjadikannya sebagai kebutuhan primer bagi setiap kegiatan, mulai dari skala kecil sampai skala besar seperti pendidikan yang melingkupi segala bidang kegiatannya. Perkembangan komputer ini juga diiringi dengan perkembangan *software* pendukungnya yaitu dalam pengolahan data dan penyajian informasi, perkembangan teknologi informasi ini berawal dari pengolahan data misalnya data nilai mahasiswa, yang mana untuk menentukan calon mahasiswa berprestasi. Pada akademi manajemen gunung leuser ini mengolah data nilai mahasiswa masih menggunakan excel, yang mana dalam menentukan calon mahasiswa berprestasi tersebut membutuhkan waktu yang cukup lama untuk mendapatkan hasil yang baik. Seiring dengan perkembangan *software* pendukungnya yaitu dalam pengolahan data dan penyajian informasi, penulis melakukan penelitian pada Akademi Manajemen Gunung Leuser. Untuk menyelesaikan permasalahan adalah dengan pemanfaatan algoritma K-Means Clustering. Algoritma K-Means adalah algoritma klustering yang paling sederhana dibanding algoritma klustering yang lain. Algoritma ini mempunyai kelebihan mudah diterapkan dan dijalankan, relatif cepat, mudah untuk diadaptasi, dan paling banyak dipraktikkan dalam tugas *Data Mining*.

2. Landasan teori

2.1 Definisi KDD

Knowledge Discovery in Database (KDD) adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam data berukuran besar. Keluaran *datamining* ini bisa dipakai untuk membantu pengambilan keputusan di masa depan.

2.2 Definisi Data Mining

Data Mining adalah proses menemukan hubungan dalam data yang tidak diketahui oleh pengguna dan menyajikannya dengan cara yang dapat dipahami sehingga hubungan tersebut dapat menjadi dasar pengambilan keputusan (Ali Akbar Yulianto dan Afia R. Fitriati (penterjemah), 2008.).

2.3 Clustering

Clustering adalah tanpa pengawasan mekanisme klasifikasi di mana satu set pola (data) , biasanya multidimensi diklasifikasikan ke dalam kelompok (*cluster*) seperti bahwa anggota satu kelompok yang sama sesuai dengan kriteria yang telah ditetapkan .*Clustering* dari set membentuk partisipasi elemen yang dipilih untuk meminimalkan beberapa ukuran perbedaan antara anggota *cluster* yang sama. *Clustering* algoritma

sering berguna dalam berbagai bidang seperti data mining , pengenalan pola , belajar teori dan lain-lain (Verma, et al 2012).

2.4 Metode K-Means

K-means merupakan suatu algoritma pengklasteran yang cukup sederhana yang mempartisi dataset kedalam beberapa kluster k. Algoritmanya cukup mudah untuk diimplementasi dan dijalankan, relatif cepat, mudah disesuaikan dan banyak digunakan (Wu & Kumar, 2009).

Diawal cara kerja sama dengan algoritma k-means, diakhir akan dilakukan perhitungan intra dan inter cluster, jika jarak intra lebih kecil dan jika jarak intra lebih besar, maka algoritma menghitung cluster baru dengan menambahkan counter k dengan satu atau k=k+1 disetiap iterasi sampai memenuhi batas validitas kualitas cluster yang berkualitas (M & Hareesha, 2012).

deviasi digunakan untuk memeriksa kedekatan titik data setiap cluster, dan dihitung sebagai:

$$d(x,y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2} \quad (1)$$

Berdasarkan data tersebut diperoleh hasil analisa kebutuhan sistem berikut:

a. Kebutuhan input

Sistem yang akan dibangun membutuhkan data input, antara lain : Data nilai mahasiswa

b. Kebutuhan Proses

Proses yang digunakan untuk mengolah data input adalah teknik *Clustering* dengan algoritma K-Means.

c. Kebutuhan Output

Output yang diharapkan adalah analisa dari hasil *Clustering* sehingga dapat diketahui informasi dari data input.

3. METODOLOGI PENELITIAN

3.1 Pendahuluan

Metodologi penelitian ini dilakukan secara sistematis yang dapat digunakan sebagai pedoman untuk peneliti dalam melaksanakan ini agar hasil yang dicapai tidak menyimpang dan tujuan yang diinginkan dapat terlaksana dengan baik dan sesuai dengan tujuan yang telah ditetapkan

3.2 Kerangka Kerja

Pada bab ini akan diuraikan kerangka kerja penelitian, kerangka kerja ini merupakan langkah-langkah yang akan dilakukan dalam penyelesaian masalah yang akan dibahas.

NO	NAMA MAHASISWA	C1	C2	C3	C4	C5	JKM1	JKM2	JKM3	JKM4	JKM5
1	ADIE HIKMAH SUKAWATI	87,4098	140,007	81,9049	116,9899	198,1706	0	0	0	0	0
2	ANANDA ARJUN HADISULAN	62,84287	110,8189	12,02776	66,96007	105,9175	0	0	1	0	0
3	DESI ARIYANTI	81,1171	110,8877	11,10833	61,47193	104,3821	0	0	1	0	0
4	DESI ERIKATRA	131,1397	11,49418	116,0848	131,9661	195,4665	0	1	0	0	0
5	ESTER ANANDA LUBIS	81,89921	100,809	14,00074	66,98803	105,0461	0	0	1	0	0
6	INDRIANI SUDHON	100,9392	198,8815	137,4118	137,3878	100,9412	0	0	0	0	1
7	IRMA YUNITA	77,05842	110,139	14,06212	67,60708	105,1915	0	0	1	0	0
8	JUNIARTI ARIYANTI	87,78863	110,5113	16,14843	61,03888	106,1009	0	0	1	0	0
9	KHARUL SULEWAN RANGKUTI	84,73989	111,1115	10,81693	61,4077	106,8069	0	0	1	0	0
10	KUSDIANI HADISULAN	130,8875	110,7272	119,4388	110,1044	199,9716	0	1	0	0	0
11	KUSDIYANI HADISULAN	81,08921	114,4168	81,08921	101,178	101,178	0	0	1	0	0
12	NOVIA SARI DAULAY	80,77958	112,814	12,71131	69,42713	106,1012	0	0	1	0	0
13	RIZKA GABRIEL HADISULAN	174,7614	186,1718	112,1716	170,1897	0	0	0	0	0	1
14	SOLINA HADISULAN	85,11781	116,1881	11,7021	101,7742	116,8746	0	0	1	0	0
15	WANGSITI HUSNA	77,86095	102,1246	12,47329	66,74069	106,8121	0	0	1	0	0
16	LETIANI RIZKA	80,78818	110,7395	11,80971	61,98811	106,7681	0	0	1	0	0
17	LIYANA ANITA WISE	81,87007	109,1283	13,0125	68,38118	101,2329	0	0	1	0	0
18	MADISULAN WIS	161,6992	174,0891	193,4477	186,7399	188,7841	0	0	0	0	1
19	MARTANI SULEWAN HADISULAN	76,08976	111,8794	14,81939	64,01766	104,1448	0	0	1	0	0
20	MARTINA YABRIANA DAULAY	81,02068	110,4872	24,4844	66,60126	107,0584	0	0	1	0	0
21	INDAH SETIYANI NIT	81,16174	114,116	14,16877	67,10919	105,1823	0	0	1	0	0
22	ANINDA SARI RITONGA	76,10961	102,1776	10,10929	61,87716	101,2041	0	0	1	0	0
23	ANANDA WIT	143,9392	109,1791	176,1161	176,3844	112,4041	0	0	0	0	1
24	MARINDA SIBERAN	81,08921	102,8968	18,01768	67,05711	101,1981	0	0	1	0	0
25	STI YANITA HADISULAN	77,94718	109,0384	19,0363	66,95877	102,5181	0	0	1	0	0
26	TURKANTI	111,931	85,80016	88,31199	117,8718	186,1221	0	1	0	0	0
27	INDA ANINDA HADISULAN	77,47923	109,8989	17,88971	68,13214	106,1168	0	0	1	0	0
28	MARTINI NIT	76,48172	102,777	10,08997	67,02475	106,8121	0	0	1	0	0
29	SARILA ANANDA NIT	76,80354	104,8189	12,04451	67,61249	104,4028	0	0	1	0	0
30	OTIS SARILAN LBS	80,71174	106,1219	14,40221	65,61116	100,0861	0	0	1	0	0
31	DELIANA ANINDA HADISULAN	78,86974	104,0961	16,01176	68,18119	100,1821	0	0	1	0	0
32	TIS SURYANDI	110,84015	84,88847	111,1018	174,7914	1	0	0	0	0	0
33	IRMA YUNITA HADISULAN	111,0094	101,1868	81,81098	170,1897	0	0	0	0	1	0
34	ANANDA WIS	85,40046	114,1164	112,8167	124,108	187,7983	1	0	0	0	0

Tabel 4.19 hasil pengclustoran

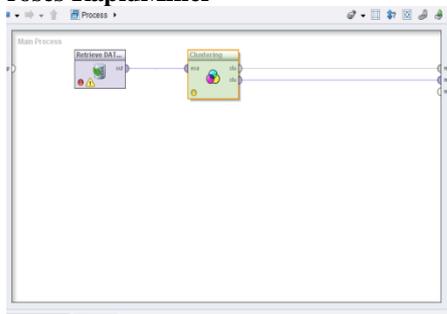
Karena iterasi 1 dan iterasi 2 dimana data tidak ada yang berpindah dan sampai disini hasil clustering sudah mencapai stabil dan konvergen

Kesimpulan :

Hasil clustering adalah

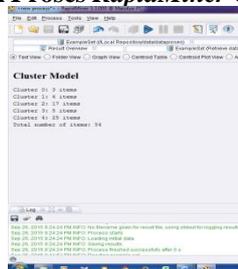
Kelompok (cluster)	Kelompok mahasiswa	Jumlah mahasiswa
1	[1;32;34]	3
2	[4;10;26]	23
3	[2;3;5;7;8;9;11;12;14;15;16;17;19;20;21;22;24;25;27;28;29;30;31]	23
4	[33]	1
5	[6;13;18;23]	4

5.1 Proses RapidMiner

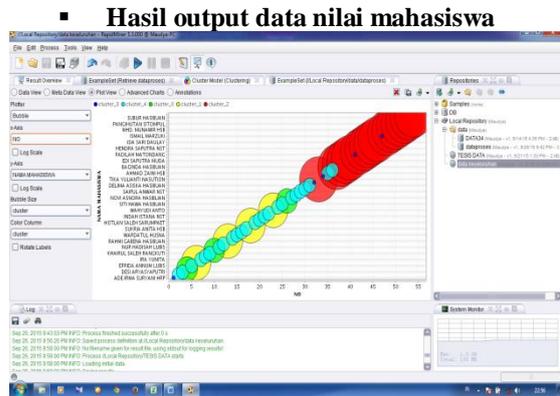


Gambar 5.12 Proses Pemindahan K-Means

5.2 Hasil Dari Proses RapidMiner Cluster Model



Gambar 5.14 Tampilan Cluster Model (Clustering) Text View



Gambar 5.18. Tampilan ExampleSet (Read Excel) Plot View

Daftar pustaka

Fadlina. (2014). "Data Mining untuk Analisa Tingkat Kejahatan Jalanan Dengan Algoritma Association Rule Metode Apriori (Studi Kasus Di Polsekta Medan Sunggal)." *Voumel.III No.1.*

GINANJAR, ANGGA MABRUR, (2012). "Penerapan Data Mining Untuk Memprediksi Kriteria Nasabah Kredit." *Edisi. I Volume. 1.*

LESTARI WIJI (2013). "Aplikasi Algoritma Competitive Network Untuk Clustering Minat Mahasiswa Terhadap Topik-Topik Penelitian." *ISSN : 2086-9436 Volume 5 Nomor 1.*

