

TINGKAT PEMAHAMAN MAHASISWA TERHADAP NORMA-NORMA AGAMA ISLAM MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING

Feri Musharyadi, M.Kom
STIKes Mercubaktijaya Padang
e-mail: ferimusharyadi@gmail.com

ABSTRACT

Pemahaman mahasiswa yang kurang terhadap norma-norma agama Islam sangat rendah dan sudah mulai memudar. Hal ini berdampak pada perilaku sehari-hari mahasiswa yang sangat jauh dari koridor tuntunan Agama Islam. Sehingga dapat menimbulkan gejala sosial di masyarakat dan menumbuhkan nilai-nilai moral yang rusak di kalangan mahasiswa. Metode *K-means* merupakan metode *clustering* yang paling sederhana dan umum. Hal ini dikarenakan *K-Means* mempunyai kemampuan mengelompokkan data dalam jumlah yang cukup besar dengan waktu komputasi yang relatif cepat dan efisien. Penulis bertujuan untuk mengkaji sejauh mana pola tingkat pemahaman mahasiswa terhadap norma-norma agama Islam.

Kata Kunci : Data Mining, K-Means, Clustering.

1. PENDAHULUAN

Tujuan utama dalam pendidikan Islam adalah agar manusia memiliki gambaran tentang Islam yang jelas, utuh dan menyeluruh. Interaksi di dalam diri manusia member pengaruh kepada penampilan, sikap, tingkah laku dan amalnya sehingga menghasilkan akhlak yang baik. (H.Moh.Solikodin Djaelani, *Peran Pendidikan Agama Islam dalam keluarga dan masyarakat, Volume 1 Nomor 2 2013*).

Sejauh ini permasalahan yang sering terjadi dalam pola pemahaman mahasiswa terhadap norma-norma agama sangat rendah dan sudah mulai memudar. Hal ini berdampak pada perilaku sehari-hari mahasiswa yang sangat jauh dari koridor tuntunan Agama Islam. Sehingga dapat menimbulkan gejala sosial di masyarakat dan menumbuhkan nilai-nilai moral yang rusak di kalangan mahasiswa.

Dengan diterapkannya norma-norma agama tersebut, mahasiswa dapat membentuk insan generasi muda, yang mana dapat membentuk kepribadian manusia dengan akhlakul karimah, serta sarat akan nilai-nilai moral Islam.

Metode *K-means* merupakan metode *clustering* yang paling sederhana dan umum. Hal ini dikarenakan *K-Means* mempunyai kemampuan mengelompokkan data dalam jumlah yang cukup besar dengan waktu komputasi yang relatif cepat dan efisien.

2. KAJIAN LITERATUR

Knowledge Discovery in Database (KDD)

Knowledge Discovery dalam Database didefinisikan sebagai ekstraksi informasi potensial, implisit dan tidak dikenal dari sekumpulan data. Data, dalam bentuk mentah, hanya kumpulan elemen, dari yang sedikit pengetahuan bisa dipetik.

Proses *knowledge discovery* melibatkan hasil dari proses *Data Mining* (proses mengekstrak kecenderungan pola suatu data), kemudian mengubah hasilnya secara akurat menjadi informasi yang mudah dipahami.

Berbagai metode tersedia untuk membantu dalam penggalian pola bahwa ketika ditafsirkan akan memberikan berharga yang mungkin sebelumnya tidak diketahui, wawasan data yang disimpan. Informasi ini dapat menjadi prediktif atau deskriptif di alam. *Data mining*, tahap ekstraksi pola KDD, dapat mengambil banyak bentuk, pilihan tergantung pada hasil yang diinginkan. KDD adalah proses multi-langkah yang memfasilitasi konversi data ke informasi yang berguna.

Data Mining

Data Mining bukanlah suatu bidang yang sama sekali baru. Salah satu kesulitan untuk mendefinisikan data mining adalah kenyataan bahwa Data Mining mewariskan banyak aspek dan teknik bidang-bidang ilmu yang sudah mapan terlebih dahulu. (Efori Buulolo, 2013).

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. *Data Mining* ini juga dikenal dengan istilah *pattern recognition* (Ong, 2013).

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 (Ong, 2013).

K-Means Cluster Analysis

Cluster Analysis merupakan salah satu metode *objek mining* yang bersifat tanpa latihan (*unsupervised analysis*), sedangkan *K-Means Cluster Analysis* merupakan salah satu metode *cluster analysis* non hirarki yang berusaha untuk mempartisi objek yang ada kedalam satu atau lebih *cluster* atau kelompok objek berdasarkan karakteristiknya, sehingga objek yang mempunyai karakteristik yang sama dikelompokkan dalam satu *cluster* yang sama dan objek yang mempunyai karakteristik yang berbeda dikelompokkan kedalam *cluster* yang lain. Tujuan *pengelompokan* adalah untuk meminimalkan *objective function* yang di set dalam proses *clustering*, yang pada dasarnya berusaha untuk meminimalkan variasi dalam satu *cluster* dan memaksimalkan variasi antar *cluster*.

K-Means merupakan algoritma *clustering* yang berulang-ulang. Algoritma *K-Means* dimulai Dengan pemilihan secara acak *K*, *K* di sini merupakan banyaknya *cluster* yang ingin dibentuk. Kemudian tetapkan nilai-nilai *K* secara random, untuk sementara nilai tersebut menjadi pusat dari *cluster* atau biasa disebut dengan *centroid*, mean atau “*means*”. Hitung jarak setiap data yang ada terhadap masing-masing *centroid* menggunakan rumus *Euclidian* hingga ditemukan jarak yang paling dekat dari setiap data dengan *centroid*. Klasifikasikan setiap data berdasarkan kedekatannya dengan *centroid*. Lakukan langkah tersebut hingga nilai *centroid* tidak berubah (stabil) (Rismawan dan Sri Kusumadewi, 2008). *K-Means Clustering* merupakan metode yang termasuk ke dalam golongan algoritma *Partitioning Clustering* (Handoyo, 2014).

3. METODE PENELITIAN

Metodologi penelitian ini merupakan langkah-langkah kerja yang perlu dilakukan agar penyusunan tesis menjadi lebih mudah dan juga dapat digunakan sebagai pedoman untuk peneliti dalam melaksanakan penelitian. Metodologi yang digunakan dalam penelitian tesis ini diperlihatkan dalam suatu kerangka kerja yang akan dijelaskan pada bab ini. Penelitian merupakan rangkaian kegiatan ilmiah dalam rangka pemecahan suatu permasalahan. Fungsi penelitian adalah mencari penjelasan dan jawaban terhadap permasalahan serta memberikan alternatif kemungkinann yang dapat digunakan untuk pemecahan masalah.

Agar langkah-langkah dalam penelitian ini dapat lebih mudah dipahami, berikut penjelasan terperinci mengenai sistematika dari kerangka kerja penelitian yang diuraikan sebagai berikut :

1. Mengidentifikasi Masalah

Pada tahap ini di lakukan pengidentifikasian masalah yang ada , dimulai dengan memperhatikan data yang telah dikumpulkan melalui pendidikan karakter khususnya mengenai norma-norma agama. Oleh Karena itu penulis ingin menganalisa serta mengelompokan informasi informasi tersebut dengan menggunakan metode k-means .

2. Menganalisa Masalah.

Dalam melakukan analisa masalah peneliti melakukan beberapa metode di antaranya yaitu metode deskriptif, dalam metode ini dilakukan pengumpulan data, kemudian disusun, dikelompokkan, dianalisa sehingga diperoleh beberapa gambaran yang jelas pada masalah penelitian. Sehingga dari analisa masalah tersebut dapat ditarik suatu kesimpulan untuk mendapatkan suatu solusi penyelesaian masalah.

3. Mempelajari Literatur.

Peneliti akan mempelajari literatur yang terkait dengan *Data Mining*, *Clustering* dan algoritma *K-Means*. Adapun literatur yang dipelajari diambil dari berbagai sumber seperti buku teks, jurnal halaman internet, *paper*, *e-book* dan karya ilmiah.

4. Mengumpulkan Data.

Untuk mendapatkan informasi yang dibutuhkan dalam penelitian, maka penulis melakukan pengumpulan data informasi melalui kuisioner berbentuk angket.

5. Menyeleksi Data.

Pada tahap ini dilakukan proses penyeleksian data dari data-data yang telah di kumpulkan, di mana data-data tersebut dilekompokan sehingga menjadi hasil pengelompokannya. Pada tahapan ini juga dilakukan tahapan *cleaning*, yaitu di mana data-data yang telah dikelompokkan untuk kemudian dilakukan pembersihan data, menghapus data duplikat, dan menghilangkan *noise*. Setelah dilakukan *cleaning*, maka dilanjutkan dengan transformasi data yang telah di bersihkan pada tahapan sebelumnya sehingga siap untuk di cluster.

6. Perancangan Algoritma *K-Means Clustering*.

Perancangan dengan penerapan algoritma *K-Means Clustering* selanjutnya sehingga menghasilkan data yang diolah hingga menghasilkan suatu pengetahuan yang baru.

7. Mengekstraksi Data (*Data Mining*) Dengan Menggunakan Algoritma *k-means*.

Pada tahap ini akan dilakukan proses perancangan dari model sistem dengan mengelompokkan data untuk menghasilkan beberapa *cluster* yang telah di tentukan dengan menggunakan algoritma *k-means*.

8. Melakukan Pengujian hasil.

Pada tahap ini akan dilakukan serangkaian pengujian untuk mevalidasi apakah sistem yang dirancang sudah sesuai dengan model yang dirancang. Adapun tahapan dalam mekanisme pengujian *Data Mining* sebagai berikut :

a) *Data Selection*

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang akan digunakan untuk proses *Data Mining*, disimpan dalam suatu berkas, terpisah dari basis data operasional.

b) *Pre-pocessing/Cleaning*

Sebelum proses *Data Mining* dapat dilaksanakan, perlu dilakukan proses *cleaning* pada data yang menjadi fokus KDD. Proses *cleaning* mencakup antara lain membuang duplikasi data, memerikasa data yang *inkonsisten*, dan memperbaiki kesalahan pada data, seperti salah cetak (*tipografi*) juga dilakukan proses *enrichment*, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan yang diperlukan untuk KDD, seperti data atau informasi eksternal.

c) *Transformation*

Coding adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses *Data Mining*. Proses *coding* dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

d) *Interpretation / Evaluation*

Pola informasi yang dihasilkan dari proses *Data Mining* perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses KDD yang disebut *interpretation*. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya.

9. Menarik Kesimpulan

Pada akhir pembahasan dilakukan proses penarikan kesimpulan yang bertujuan untuk membandingkan hasil yang diperoleh dari tahap implementasi sistem yang dibuat secara manual.

4. HASIL DAN PEMBAHASAN

Di dalam *Data Mining* terdapat beberapa teknik pengolahan data agar data yang di olah lebih bermanfaat dan bernilai ilmu. Salah satu metode yang terdapat dalam *Data Mining* adalah *K-Means*. Teknik atau metode dari analisis *k-means* dapat dijelaskan proses algoritma *k-means* untuk menyelesaikan masalah, yaitu sebagai berikut :

Pengumpulan Data

Metode pengumpulan data yang digunakan adalah dengan wawancara langsung dengan dosen pendidikan karakter pada STIKes MERCUBAKTIJAYA Padang. Berikut ini adalah beberapa data rekapitulasi pelaksanaan pendidikan karakter STIKes MERCUBAKTIJAYA Padang :

NO	NAMA	NIM	PK	BA	SOS	SJ
1	NOFRIATI UTAMI PUTRI	14111889	27	7	5	5
2	ROSA MASFIKA	14111890	33	5	6	4
3	NADIA ERWIN	14111891	29	6	6	5
4	CYNTHIA LAURA DIANA	14111892	28	6	7	6
5	ANGGUN RAFIKA	14111893	38	7	7	6
6	AJENG PUTRI MEREYSA	14111894	30	8	7	8
7	DESY YULIANTI	14111895	29	8	8	8
8	FAIZAH HAZNI	14111897	31	8	5	5
9	DEBY CINTHYA TRI MANDA	14111899	31	3	8	4
10	ARIFA PERTIWI	14111900	35	5	5	4
11	OLGA VONINDA	14111901	36	6	5	5
12	OLYVIA DELFITA	14111902	32	5	3	3
13	CERAH MUTIARA HATI	14111903	36	6	5	7
14	AHMAD MUSOHUR SIREGAR	141211000	39	7	4	4
15	MIFTAHUL MUBARAK	141211009	30	5	3	6
16	MUHAMMAD ALZAT	141211013	35	7	3	3
17	YUDA PRATAMA PUTRA	14121973	35	7	3	3
18	JERI SANDRA	14121974	39	7	6	5
19	MUTIA MONICA	141211001	32	7	4	5
20	YOHANA MONICA ARIVITA	141211002	28	5	5	7
21	PUTRI LIDIA SISKA	141211003	30	7	6	2
22	SAKRIAH HIRLIANI	141211004	38	7	5	8
23	MURHAFIDDAH YULIA SARI	141211005	30	7	6	4
24	WADIA YUNISDWI PUTRI	141211006	38	8	6	4
25	MBA KHAIRUNNISA H	141211007	36	8	6	4
26	MUHYUR FANI	141211008	31	7	5	6
27	ATIQAH	141211010	36	8	6	4
28	ATIKA PUTRI	141211012	35	6	5	4
29	NELA INDRIANI	141211013	31	7	6	4

Menentukan jumlah Cluster

Sebagaimana telah dijelaskan pada sub bab sebelumnya bahwa untuk pengelompokan data pada pengujian yang pertama ini adalah sebanyak 3 cluster, sehingga dapat ditentukan untuk nilai k adalah k = 3. Dimana nantinya berkemungkinan untuk dapat di artikan Cluster 1 terdiri dari mahasiswa yang Memahami akan norma norma agama, Cluster 2 terdiri dari mahasiswa yang kurang memahami dan Cluster 3 terdiri dari mahasiswa yang tidak memahami akan norma norma agama atau sebaliknya.

Menentukan titik pusat Cluster

Penentuan centroid awal dilakukan secara acak dari data/objek yang tersedia sebanyak jumlah cluster k. Kemudian untuk menghitung centroid pada cluster berikutnya ke-i, digunakan rumus sebagai berikut :

$$V = \frac{\sum_{i=1}^n X_i}{n}; i = 1,2,3 \dots n$$

C1	NOFRIATI UTAMI PUTRI	27	7	5	5
C2	DEBY CINTHYA TRI MANDA	31	5	6	4
C3	OLGA VONINDA	36	6	5	5

Hitung jarak setiap data yang ada terhadap setiap pusat Cluster

Untuk menghitung jarak setiap data yang ada terhadap pusat cluster yaitu dengan menggunakan *Euclidean Distance* untuk melakukan perhitungan jarak setiap data terhadap titik pusat cluster

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

Untuk C1 :

$$d(x,y) = \sqrt{(27 - 27)^2 + (7 - 7)^2 + (5 - 5)^2 + (5 - 5)^2}$$

$$= 0$$

$$d(x,y) = \sqrt{(27 - 33)^2 + (7 - 5)^2 + (5 - 6)^2 + (5 - 4)^2}$$

$$= 6,48$$

$$d(x,y) = \sqrt{(27 - 29)^2 + (6 - 7)^2 + (6 - 5)^2 + (5 - 5)^2}$$

$$= 2,45$$

$$d(x,y) = \sqrt{(27 - 28)^2 + (6 - 7)^2 + (7 - 5)^2 + (6 - 5)^2}$$

$$= 2,65$$

$$d(x,y) = \sqrt{(27 - 38)^2 + (7 - 7)^2 + (7 - 5)^2 + (6 - 5)^2}$$

$$= 11,22$$

Sehingga dapat kita buat bentuk iterasi pertama dari cluster :

Tabel 4.3 Posisi Cluster Pada Iterasi Pertama

ITERASI 1					
C1	C2	C3	C1	C2	C3
0	4,690416	9,055385	1	0	0
6,480741	2	3,464102	0	1	0
1,732051	3,316625	8,062258	1	0	0
2,645751	3,872983	8,306624	1	0	0
11,22497	7,615773	3,162278	0	0	1
4,795832	4,358899	7	0	1	0
3,872983	4,582576	7,681146	1	0	0
4,123106	1,732051	5	0	1	0
4,690416	0	5,291503	0	1	0
8,306624	4,123106	1,732051	0	0	1
9,055385	5,291503	0	0	0	1
6,082763	3,316625	5	0	1	0
9,273618	6	2	0	0	1
12,08305	8,485281	3,464102	0	0	1
4,690416	0	5,291503	0	1	0
8	4,690416	1,414214	0	0	1
8	4,690416	1,414214	0	0	1
12,04159	8,306624	3,316625	0	0	1
5,09902	3,162278	4,242641	0	1	0
3	4,358899	8,306624	1	0	0
			5	7	8

Perhitungan pusat *cluster* baru :

$$C11 = (27+28+28+29+28)/5 = 28$$

$$C12 = (7+6+6+6+5)/5 = 6$$

$$C13 = (5+6+7+6+5)/5 = 5,8$$

$$C14 = (5+5+6+8+7)/5 = 6,2$$

$$C21 = (33+30+31+31+32+31+32)/7 = 31,42$$

$$C22 = (5+6+6+5+5+5+7)/7 = 5,57$$

$$C23 = (6+7+5+6+3+6+4)/7 = 5,28$$

$$C24 = (4+8+5+4+3+4+5)/7 = 4,71$$

$$C31 = (38+35+36+36+39+35+35+39)/8$$

$$= 36,62$$

$$C32 = (7+5+6+6+7+7+7+7)/8 = 6,5$$

$$C33 = (7+5+5+5+4+5+5+6)/8 = 5,25$$

$$C34 = (6+4+5+7+4+5+5+5)/8 = 5,125$$

C1	28	6	5,8	6,2
C2	31,42857	5,571429	5,285714	4,714286
C3	36,625	6,5	5,25	5,125

Ulangi dari langkah ketiga hingga nilai dua iterasi tidak lagi berubah

Karena proses yang dilakukan baru pada pengulangan ke 1, maka perlu dilakukan beberapa pengulangan lagi untuk dapat membandingkan nilai dari dua iterasi terakhir. Jika nilai dua iterasi terakhir tersebut telah sama, maka proses iterasi telah selesai, dan jika tidak maka ulangi lagi langkah ke 3.

Hasil iterasi Berdasarkan hasil pengelompokkan data menggunakan metode *k-means clustering*, didapatkan hasil *clustering* hingga pengulangan ke 3 dan pada pengulangan ke 3 nilai dari iterasi tidak lagi mengalami perubahan, sehingga proses pengulangan dihentikan pada pengulangan ke 3. Berikut adalah hasil iterasi yang telah dilakukan.

Tabel 4.5 Posisi Cluster Pada Iterasi Ketiga

ITERASI 3						
C1	C2	C3	C1	C2	C3	
2,455153	4,972145	9,64203	1	0	0	
5,387743	1,748015	4,149548	0	1	0	
1,536591	3,922867	8,672874	1	0	0	
1,166667	4,588633	8,858259	1	0	0	
9,782354	7,051399	2,443103	0	0	1	
2,455153	4,66071	7,447735	1	0	0	
1,641476	4,80162	8,198704	1	0	0	
3,218868	1,178511	5,654091	0	1	0	
3,789606	1,312335	5,976517	0	1	0	
7,259094	3,374743	2,493742	0	0	1	
7,875772	4,440971	0,847791	0	0	1	
5,974483	2,392117	5,763571	0	1	0	
7,74776	5,201496	2,05396	0	0	1	
11,18158	7,553513	2,952753	0	0	1	
3,789606	1,312335	5,976517	0	1	0	
6,97814	3,749074	1,723006	0	0	1	
6,97814	3,749074	1,723006	0	0	1	
10,81794	7,597514	2,543374	0	0	1	
4,549115	2,013841	4,818584	0	1	0	
1,536591	4,66071	8,956492	1	0	0	
			6	6	8	

5. SIMPULAN

- Berdasarkan proses pengulangan 2 dan 3, terlihat bahwa titik pusat *cluster* yang diperoleh tetap sama, dan tidak ada lagi data yang berpindah dari satu *cluster* ke *cluster* yang lain, sehingga proses iterasi dihentikan pada pengulangan ke 3.
- Dapat disimpulkan bahwa tingkat pemahaman mahasiswa terhadap norma-norma agama Islam sesuai kelompoknya sebagai berikut :
Kelompok 1 : 1,3,4,6,7,20
Kelompok 2 : 2,8,9,12,15,19
Kelompok 3 : 5,10,11,13,14,16,17,18
- Penggunaan *Data Mining* Sangat membantu dalam menggali informasi tingkat pemahaman mahasiswa terhadap norma-norma agama Islam.

6. REFERENSI

- Ahmad R. 2014. *ClusteringData Cuaca Untuk Pengenalan Pola Perioditas Iklim Wilayah Pelabuhan Dengan Metode Fuzzy C-Means*. Vol.3 1Juni 2014.
- Tahta A, Budi S. 2012. *Analisa Perbandingan Metode Hierarchical Clustering, K-Means dan Gabungan Keduanya dalam Cluster Data*. Vol.1, Sept 2012

- Tedy R, Sri K 2008. *Aplikasi K-Means Untuk Pengelompokkan Mahasiswa Berdasarkan Nilai Body Mass Index (BMI) & Ukuran Rangka*.
- R Handoyo, R Rumani. 2014. *Perbandingan Metode Clustering Menggunakan Metode Single Linkage dan K-Means pada pengelompokan Dokumen*. Vol.15, No2, Oktober 2014.
- Ong, 2013. *Implementasi Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing Presiden University*. Vol.12, No.1, Juni 2013.
- Lindawati, 2008. *Data Mining Dengan Teknik Clustering Dalam Pengklasifikasian Data Mahasiswa Studi Kasus Prediksi Lama Studi Mahasiswa Universitas Bina Nusantara*.
- Efori Buulolo, 2013. *Implementasi Algoritma Apriori Pada Sistem Persediaan Obat (Studi Kasus : Apotik Rumah Sakit Estomihi Medan)*. Vol.IV, No.1, Agustus 2013
- Ediyanto, 2013. *Pengklasifikasian Karakteristik Dengan Metode K-Means Clustering Analisis*. Vol.2, No.2, 2013