

PENENTUAN SEKOLAH TERDEKAT UNTUK VISITASI ASESOR MENGGUNAKAN METODE ALGORITMA *K-MEANS* BERBASIS WEB

Zainal Ikhwan Muhammad

Teknik Informatika, Fakultas Teknologi Informasi, Universitas Hasyim Asy'ari

Zainal.studio@gmail.com

IGL. Putra Eka Prisma

Teknik Informatika, Fakultas Teknologi Informasi, Universitas Hasyim Asy'ari

ekaprismana@gmail.com

Chamdan Mashuri

Teknik Informatika, Fakultas Teknologi Informasi, Universitas Hasyim Asy'ari

chamdanmashuri@unhasy.ac.id

Abstrak

Penentuan sekolah terdekat untuk visitasi asesor merupakan suatu langkah yang dibuat guna mempermudah para asesor untuk melaksanakan akreditasi sekolah. Akreditasi sekolah merupakan upaya dalam peningkatan mutu pendidikan nasional. Badan Akreditasi Provinsi Sekolah atau Madrasah (BAP- S/M) melakukan perekrutan asesor dan dibagi tugasannya diberbagai tempat atau wilayah untuk melaksanakan visitasi. Penelitian ini bertujuan untuk merancang sistem penentuan sekolah terdekat berbasis *website* dan menerapkan metode *k-means* untuk sistem tersebut. Metode ini digunakan untuk mengelompokkan data dengan menentukan jumlah *cluster* atau kelompok terdahulu, menghitung pusat *centroid* dan mengelompokkan data yang memiliki kesamaan variabel. Perhitungan metode ini menghasilkan beberapa iterasi yang memiliki nilai *cluster*. Dari beberapa iterasi tersebut digunakan nilai *cluster* yang paling minimal untuk menentukan kelompok sekolah dalam satu provinsi. Hasil penelitian ini adalah sebuah sistem informasi penentuan sekolah terdekat untuk visitasi asesor. Pengujian sistem ini dilakukan pada sekolah di wilayah provinsi Jawa Timur dengan titik koordinat sebagai variabel yaitu koordinat *latitude* dan *longitude*. Dari hasil pengujian dengan 20 data data sekolah ke dalam 3 cluster, diperoleh hasil cluster 1 dengan titik pusat koordinat (-7,213605, 112,769658) berjumlah 8 sekolah, cluster 2 dengan titik pusat koordinat (-7,202459, 112,636323) berjumlah 7 sekolah dan cluster 3 dengan titik pusat koordinat (-7,249299, 112,636909) berjumlah 5 sekolah.

Kata Kunci: *K-Means*, *Clustering*, Akreditasi, Sekolah, Web.

Abstract

The closest school determination to visitation of assessors is a step made to facilitate the assessors to implement school accreditation. School Accreditation is an effort to improve the quality of National Education. The provincial school accreditation body or Madrasah (BAP-S/M) conducts recruitment of assessors and is divided by a number of places or regions to perform visitation. The Research aims to design a nearby school-based *website* determination System and implement the *K-means* method for such systems. This method is used to group the data by specifying the number of *clusters* or previous groups, calculating the *centroid* Center and grouping the data that has the similarity of variables. This method calculation generates multiple iterations that have *cluster* values. Of these iterations used the least number of *cluster* values to determine the group of schools within one province. The result of this research is a nearby school determination information system for the visitation of assessors. The testing of this system was conducted at the school in the province of East Java with the coordinate point as a variable of *latitude* and *longitude* coordinates. From the test results with 20 school data data into 3 clusters, obtained the result of cluster 1 with the coordinate center point (-7,213605, 112,769658) amounting to 8 schools, cluster 2 with the coordinate center point (-7,202459, 112,636323) amounting to 7 school and cluster 3 with coordinate center point (-7,249299, 112,636909) amounting to 5 schools.

Keywords : *K-Means*, *Clustering*, Accreditation, School, Web.

PENDAHULUAN

Akreditasi lembaga merupakan proses penilaian secara menyeluruh terhadap tolak ukur dan program pendidikan, hal ini bertujuan untuk memberikan pengakuan kepada pencapaian standar mutu pendidikan yang telah ditetapkan dengan pemberian sertifikat status akreditasi. Materi penilaian untuk akreditasi lembaga mengambil komponen standar pendidikan (Menteri Pendidikan Nasional, 2007). Untuk akreditasi, Badan Akreditasi Provinsi Lembaga atau Sekolah (BAP-L/S) melaksanakan pengambilan asesor. Para asesor akan dibagi tugasannya diberbagai tempat atau wilayah. Metode *k-means clustering* dapat menentukan kelompok wilayah sekolah menggunakan data jarak antar sekolah.

Telah dilakukan penelitian menentukan strategi promosi STIKOM Dinamika Bangsa Jambi menggunakan Metode *Clustering K-Means*. Pada penelitian ini melakukan *clustering* data alumni dengan tujuan untuk mengetahui karakteristik profil alumni pada STIKOM Dinamika Bangsa Jambi Kupang dengan menggunakan metode *k-means*. Atribut yang digunakan dalam proses *clustering* yaitu Mama Mahasiswa, Program Studi, IPK dan Asal Sekolah. Karakteristik alumni tiga prodi (Teknik Informatika, Sistem Komputer dan Sistem Informasi). Hasil yang didapatkan dalam pengujian penelitian yaitu ada 3 pengelompokan *clustering* dapat digunakan untuk merekomendasi strategi yang tepat kepada tim promosi di daerah provinsi Jambi agar lebih efektif dan efisien (Jasmir, 2017).

Telah dilakukan penelitian penentuan kualitas *spareepart* mobil pada bengkel andesco menggunakan metode *K-Means* berbasis web. Penulis dalam penelitian ini melakukan *clustering* terhadap data *spareepart* bengkel Andesco menggunakan metode *k-means clustering*, dengan tujuan untuk mengetahui tingkat kelarisan *spareepart* dari penjualan tahun ke tahun.. Hasil yang didapatkan dalam pengujian yaitu ada beberapa pengelompokan *spareepart* yang sering laku atau terjual serta dapat melihat informasi produk mana saja yang banyak diminati dan yang kurang diminati (Putra, 2018).

Adapun tujuan penelitian ini yakni untuk menentukan lembaga terdekat untuk visitasi asesor menggunakan metode *k-means clustering*.

DASAR TEORI

Akreditasi Sekolah

Akreditasi merupakan aktivitas penilaian mutu akademik berdasarkan kriteria yang ditetapkan. Aktivitas akreditasi akademik sekolah dilaksanakan oleh pemerintah atau Lembaga mandiri yang mempunyai wewenang untuk menilai kelayakan program pendidikan yang bersifat formal atau non formal pada setiap jenjang dan jenis pendidikan, berlandaskan kriteria yang sudah ditetapkan, sebagai bentuk akuntabilitas global yang telah dilaksanakan secara objektif, adil dan menyeluruh dengan menerapkan aspek dan kriteria sesuai dengan standar nasional pendidikan (Awaludin, 2017).

Clustering

Clustering merupakan sebuah prosedur untuk mencari data yang mempunyai kesamaan karakteristik pada data tersebut. *Clustering* adalah salah satu metode data mining yang bersifat *unsupervised* yaitu prosesnya tanpa adanya latihan data dan tidak memerlukan target keluaran. Dalam data mining ada 2 jenis metode *clustering* yang dilakukan dalam proses pengelompokan data yakni *hierarchical clustering* dan *non-hierarchical clustering* (Ong, 2013).

K-Means

K-Means merupakan salah satu algoritma pada data mining yang dapat dipakai untuk proses pengelompokan data dan berupaya mempartisi data menjadi beberapa kelompok *cluster* sehingga data yang mempunyai karakteristik sama nanti akan dikelompokkan ke dalam *cluster* yang sama dan data lainnya akan dikelompokkan ke dalam data yang sesuai dengan kesamaan karakteristik data tersebut (Putri, 2015).

Pengelompokan data menggunakan *K-Means* dapat dilakukan proses sebagai berikut (Prasetyo, 2012) :

1. Menentukan jumlah kelompok.
2. Alokasikan data ke dalam kelompok secara acak.
3. Menghitung titik pusat *centroid* dari data yang telah ada. *Centroid* diambil dari rata-rata nilai. Jika M memunculkan jumlah data dalam sebuah kelompok, maka i menyatakan fitur ke- i dalam sebuah kelompok, kemudian p menguraikan dimensi data, maka persamaan untuk menghitung *centroid* fitur ke- i menggunakan persamaan 1.

$$C_i = \frac{1}{M} \sum_{i=1}^M x_i \quad (2.1)$$

persamaan (1) dilakukan sebanyak p dimensi dari $i = 1$ sampai dengan $i = p$.

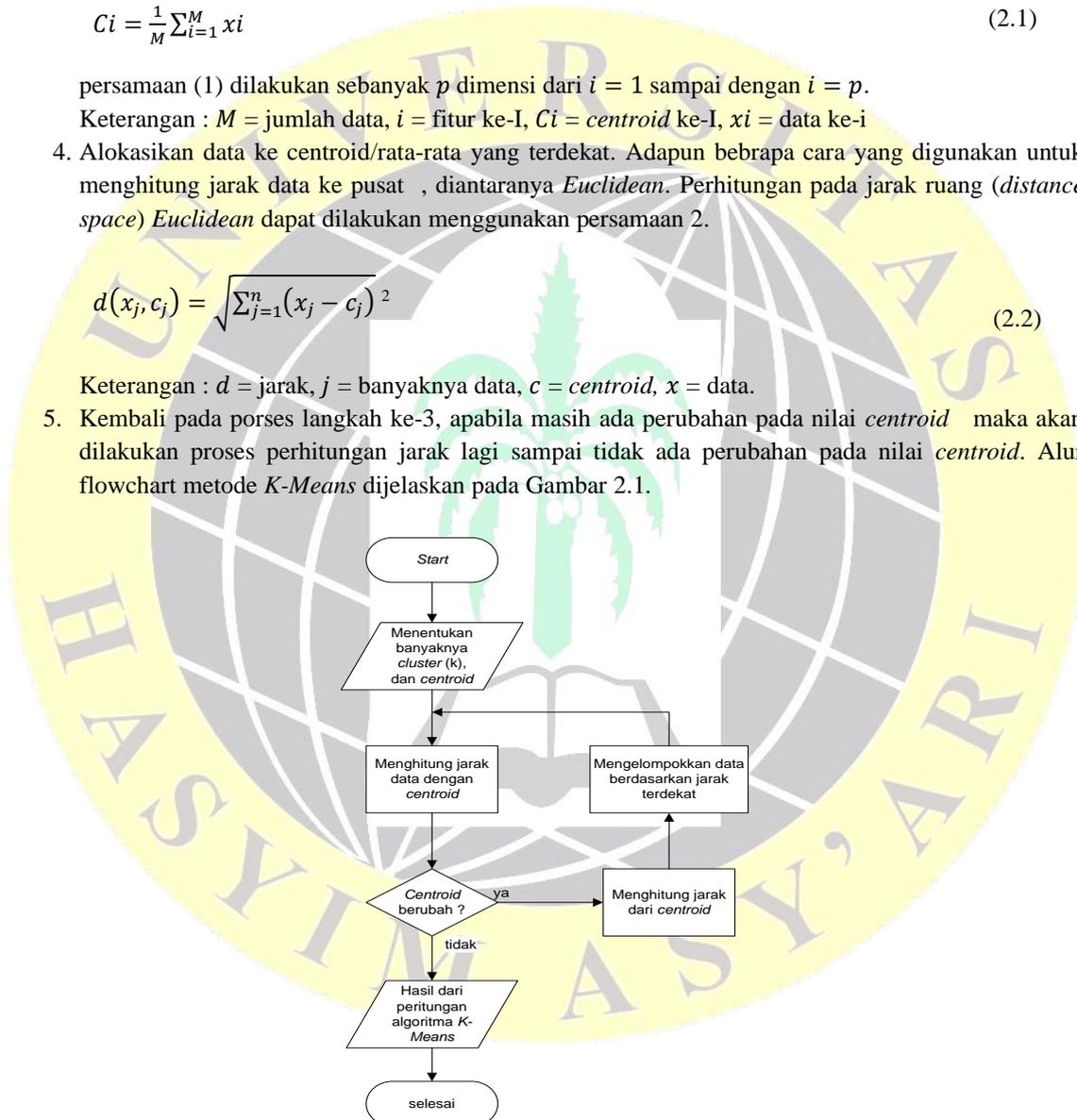
Keterangan : M = jumlah data, i = fitur ke- i , C_i = *centroid* ke- i , x_i = data ke- i

4. Alokasikan data ke centroid/rata-rata yang terdekat. Adapun beberapa cara yang digunakan untuk menghitung jarak data ke pusat, diantaranya *Euclidean*. Perhitungan pada jarak ruang (*distance space*) *Euclidean* dapat dilakukan menggunakan persamaan 2.

$$d(x_j, c_j) = \sqrt{\sum_{j=1}^n (x_j - c_j)^2} \quad (2.2)$$

Keterangan : d = jarak, j = banyaknya data, c = *centroid*, x = data.

5. Kembali pada proses langkah ke-3, apabila masih ada perubahan pada nilai *centroid* maka akan dilakukan proses perhitungan jarak lagi sampai tidak ada perubahan pada nilai *centroid*. Alur flowchart metode *K-Means* dijelaskan pada Gambar 2.1.



Gambar 2.1 Flowchart algoritma *K-Means* (Prasetyo, 2012)

PEMBAHASAN

Pre-processing Data

Pre-processing data adalah langkah tahapan dari data mining yaitu suatu proses yang dilakukan untuk mempersiapkan data mentah menjadi data siap untuk dilakukan proses selanjutnya..

Tahapan *pre-processing* data ini adalah tahapan yang sangat penting karena dalam proses data mining membutuhkan data yang konsisten dalam penulisannya, tidak ada data yang kosong, dan duplikasi data. Dalam penelitian ini tahapan *pre-processing* data memiliki dua proses yaitu:

Penghitungan Data Awal

Penghitungan data awal merupakan tahapan *pre-processing* data yang pertama dengan mencari data koordinat *latitude* dan *longitude* dari alamat data sekolah yang ada di Jawa Timur.

Penghitungan Data Siap

Penghitungan data siap adalah tahapan *pre-processing* data yang kedua. Pada tahap ini akan mengambil hasil dari data awal yang sudah diproses, dari data hasil proses *presentase* ini merupakan data yang sudah siap untuk dilakukan proses *k-means clustering*. Berikut data siap dapat dilihat pada tabel 4.10:

Tabel 3.1 Penghitungan data awal

ID	Nama Sekolah	Latitude	Longitude
1	MIS Ibnu Hasyim	-7,249180	112,708415
2	MIS NU MIFTAHUL ULUM	-7,229174	112,683689
3	MIS TAMAN JAYA	-7,229286	112,689527
4	MIS TARBIYATUS SHIBYAN	-7,249401	112,706553
5	MIS AL FATICH	-7,210207	112,652991
6	MIS MIFTAHUR RAHMAN	-7,249299	112,636909
7	MIS SULTAN AGUNG	-7,202459	112,636323
8	SD ISLAM AL MUAWANAH	-7,249421	112,641153
9	SD NEGERI SEMEMI II	-7,252401	112,645101
10	MIS AL HIDAYAH	-7,237418	112,714337
11	MIS AL HUDA	-7,284341	112,751587
12	MIS AL JIHAD	-7,285113	112,749664
13	SDN MOJO I 220	-7,265998	112,767550
14	MIS SAFINDA	-7,333995	112,765995
15	MIS EL RAHMAH PLUS QIRAATI	-7,337946	112,710956
16	MIS MASJID AL AKBAR	-7,336398	112,715243
17	MIS AL - ICHSAN	-7,212323	112,769819
18	MIS USWATUN HASANAH	-7,213605	112,769658
19	MIS YATABU	-7,232421	112,760844
20	MIS BINA BANGSA	-7,238014	112,727308

Proses Perhitungan K-Means Clustering

Penghitungan pada metode *K-Means* adalah untuk menghitung data dari *cluster* yang sudah di tentukan sesuai dengan jarak tertentu. Dalam proses perhitungan *K-Means* dibutuhkan data jumlah kelompok, data sekolah dan data koordinat *latitude* *longitude*. Proses perhitungan *K-Means* mempunyai tahapan sebagai berikut :

Menentukan Kelompok (Cluster)

Setelah melakukan tahapan *pre-processing*, selanjutnya adalah menentukan jumlah *cluster* yang akan dibentuk dan menentukan titik pusat *cluster* pertama secara random.. Penentuan pusat *cluster* awal dapat dilihat pada Tabel 4.2 berikut:

Tabel 3.2 Titik Pusat *Cluster*

Cluster	Nama	Latitude	Longitude
Cluster 1	MIS USWATUN HASANAH	-7,213605	112,769658
Cluster 2	MIS SULTAN AGUNG	-7,202459	112,636323
Cluster 3	MIS MIFTAHUR RAHMAN	-7,249299	112,636909

Menghitung jarak data dengan *K-Means*

Setelah menentukan nilai titik pusat setiap *cluster*, proses selanjutnya adalah menghitung jarak data ke setiap pusat *cluster*. Penghitungan yang akan diuji pada Sistem ini adalah proses penghitungan *clustering*. Disini peneliti akan menggunakan data *latitude* dan *longitude*.

Untuk proses perhitungan metode *K-Means Clustering* menggunakan rumus 2.2. Sebagai berikut :

Data ke-1 : C1 (-7,213605, 112,769658)

$$= \sqrt{\frac{((-7,249180) - (-7,213605))^2}{(112,708415 - 112,769658)^2}} = 0.0013$$

Data ke-1 : C2 -7,202459, 112,636323)

$$= \sqrt{\frac{((-7,249180) - (-7,202459))^2}{(112,708415 - 112,636323)^2}} = 0.1339$$

Data ke-1 : C3 (-7,249299, 112,636909)

$$= \sqrt{\frac{((-7,249180) - (-7,249299))^2}{(112,708415 - 112,636909)^2}} = 0.1380$$

Berikut hasil dari penghitungan pertama dari jarak data ke setiap *cluster* dapat dilihat pada Tabel 4.13:

Tabel 3.3 Hasil Perhitungan rumus *ecludien*

No	Jarak Ke-		
	Cluster 1	Cluster 2	Cluster 3
1	0.0013	0,929861111	0,958333333
2	0,810416667	0,127777778	0,29375
3	0,418055556	0,59375	0,54375
4	0,506944444	0,981944444	0,832638889
5	0,515972222	0,974305556	0,821527778
6	0,339583333	0,678472222	0,632638889
7	0,954861111	1,074305556	0,802083333
8	0,491666667	0,596527778	0,496527778
9	0,932638889	1,079861111	0,813194444
10	0,954861111	0,325	0.0000
11	0,606944444	0,377777778	0,353472222

12	0,836111111	1,282638889	1,072222222
13	0,929166667	0.0000	0,325
14	0,567361111	0,413888889	0,390972222
15	0,504166667	0,586805556	0,483333333
16	0.0000	0,929166667	0,954861111
17	0,144444444	0,889583333	0,86875
18	0,926388889	0,327777778	0.0042
19	0,363888889	1,0125	0,914583333
20	0,90625	0,352083333	0.0088

Dari tabel diatas dapat diketahui hasil proses iterasi awal. selanjutnya dilakukan proses penentuan pusat *cluster* baru untuk melakukan proses iterasi selanjutnya. Proses tersebut akan terulang sampai pada cluster dimana hasil perhitungannya tidak berubah. Jadi itulah hasil akhirnya.

Setelah dilakukan beberapa kali proses iterasi maka dapat diketahui dari hasil anggota pada setiap *cluster*. Berikut adalah hasil akhir dari keanggotaan dari setiap *cluster*.

Tabel 3.4 Hasil perhitungan *clustering*

Cluster	Nama Sekolah
Cluster 1	MIS AL - ICHSAN
Cluster 1	MIS AL HIDAYAH
Cluster 1	MIS BINA BANGSA
Cluster 1	MIS Ibnu Hasyim
Cluster 1	MIS TARBIYATUS SHIBYAN
Cluster 1	MIS USWATUN HASANAH
Cluster 1	MIS YATABU
Cluster 1	SD NEGERI MOJO I 220
Cluster 2	MIS AL FATICH
Cluster 2	MIS MIFTAHUR RAHMAN
Cluster 2	MIS NU MIFTAHUL ULUM
Cluster 2	MIS SULTAN AGUNG
Cluster 2	MIS TAMAN JAYA
Cluster 2	SD ISLAM AL MUAWANAH
Cluster 2	SD NEGERI SEMEMI II
Cluster 3	MIS AL HUDA
Cluster 3	MIS AL JIHAD
Cluster 3	MIS EL RAHMAH PLUS QIRAATI
Cluster 3	MIS MASJID AL AKBAR
Cluster 3	MIS SAFINDA

Menurut hasil akhir yang terlihat pada tabel diatas maka dapat dilihat bahwa dari 20 sekolah yang berada pada provinsi Jawa Timur terdapat 8 sekolah pada cluster 1 atau kelompok 1, 7 sekolah yang berada pada cluster 2 atau kelompok 2 dan 5 sekolah yang berada pada cluster 3 atau kelompok 3.

KESIMPULAN

Berdasarkan pembahasan hasil penelitian yang telah dilakukan, algoritma K-Means mampu mengelompokkan sekolah tetdekak untuk visitasi asesor ke dalam 3 cluster, yang mana cluster 1 adalah kelompok asesor pertama, cluster 2 adalah kelompok asesor kedua, cluster 3 adalah kelompok asesor ketiga. Dari hasil pengujian 20 data sekolah dihasilkan sebanyak 8 sekolah untuk kelompok asesor pertama, 7 sekolah untuk kelompok asesor kedua dan 5 sekolah untuk kelompok asesor ketiga

DAFTAR PUSTAKA

- Awaludin, Aulia Ar Rakhman. 2017. "Akreditasi Sekolah sebagai Suatu Upaya Penjaminan Mutu Pendidikan di Indonesia". *Jurnal SAP*. Vol.2 No.1.
- Badan Akreditasi Nasional Sekolah/Madrasah. 2010. *Pokok-Pokok Kebijakan dan Pedoman Akreditasi Sekolah /Madrasah*. Jakarta: Badan Akreditasi Nasional Sekolah/Madrasah (BAN-S/M).
- Jasmir. 2017. "Analisis Profil Akademik Alumni Dengan Menggunakan Metode Klasterisasi K-Means Untuk menentukan Strategi Promosi STIKOM Dinamika Bangsa Jambi". *Media Sisfo*. Vol.11 No.1.
- Menteri Pendidikan Nasional. 2005. Peraturan pemerintah republik indonesia nomor 19 tahun 2005 tentang standar nasional pendidikan.
- Ong, Johan Oscar. 2013. "Jurnal Implementasi Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing President University". *Jurnal Ilmiah Teknik Industri*, 12(1):10–20.
- Prasetyo, E., 2012. *Data Mining : Konsep dan Aplikasi Menggunakan MATLAB*. Yogyakarta : Andi.
- Putra, Teri Ade. 2018. "Perancangan Data Mining Untuk Menentukan Tingkat Kelarisan Sparepart Mobil Pada Bengkel Andesco Menggunakan Metode Clustering Dengan Algoritma K-Means Berbasis Web". *JITI*. Vol.1 No
- Putri, Dewi Eka. 2015. "Metode Non Hierarchy Algoritma K-Means Dalam Mengelompokkan Tingkat Kelarisan Barang (Studi Kasus : Koperasi Keluarga Besar Semen Padang)". *SENATKOM* ISSN: 2460-4690.