

# SISTEM PENDUKUNG KEPUTUSAN PENJURUSAN SISWA MAN 1 JEMBER MENGUNAKAN ALGORITMA K-MEANS CLUSTERING

<sup>1</sup> *Risqy Dwi Azhary (1110651065)*

<sup>2</sup> *Daryanto, S.Kom, M.Kom,* <sup>3</sup> *Henny Wahyu Sulisty, S.Kom*

*Jurusan Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jember*

## ABSTRACT

MAN is middle education stairs that give priority to prepare students for continue their more stairs in specialization. The specialization form is make major, its start at 10 grade, the major is about mathematic and science, social, Language, and Religion. So, majoring is a strategic way to give facilities for students to distribute their talents, proclivities, and capabilities on optimum manner. Therefore, school must be able to major the students in a certain place. To avoid possibility mistake major of student to choose, need to do measuring and estimating efectivity of the majoring plan. For it in grouping students of X grade MAN 1 Jember 374 students in the majores, than use K-Means Clustering methode to help determine criterion major in MAN 1 Jember. Trial was do about 100 datas be found 69 similarity and 31 diverification. So, the percentage error level is about 44,9%.

**Key Words** : *Decision Support System Major Students Of MAN 1 JEMBER Using The K-Means Clustering Algorithm.*

## ABSTRAK

Madrasah Aliyah Negeri (MAN) merupakan jenjang pendidikan menengah yang mengutamakan penyiapan siswa untuk melanjutkan pendidikan ke jenjang yang lebih tinggi dengan cara pengkhususan. Perwujudan pengkhususan tersebut berupa diselenggarakannya penjurusan yang dimulai di kelas X (sepuluh), yakni penjurusan Matematika dan Ilmu Pengetahuan Alam, Ilmu-Ilmu Sosial, Bahasa, dan Agama. Dengan demikian, penjurusan merupakan upaya strategis dalam memberikan fasilitas kepada siswa untuk menyalurkan bakat, minat, dan kemampuan yang dimilikinya untuk dikembangkan secara optimal. Oleh karena itu, sekolah harus mampu menempatkan siswa ke dalam jurusan yang tepat. Untuk menghindari kemungkinan terjadinya ketidak tepatan penjurusan siswa dalam pilihan jurusan tertentu, perlu adanya upaya pengukuran dan penilaian keefektifan perencanaan penjurusan, keefektifan pelaksanaan penjurusan, keberhasilan siswa setelah penjurusan, serta kendala-kendala yang dihadapi dalam penjurusan. Kemudian algoritma *k-means* adalah metode klasterisasi yang paling terkenal dan banyak digunakan di berbagai bidang karena sederhana, mudah diimplementasikan, serta memiliki kemampuan untuk mengklaster data yang besar, mampu menangani data banyak.

**Kata kunci** : *Sistem Analisa Penentuan Penjurusan Siswa MAN 1 Jember Menggunakan Algoritma K-Means Clusteri*

## BAB I PENDAHULUAN

### 1.1 Latar Belakang

Madrasah AliyahNegeri (MAN) merupakan jejang pendidikan menengah yang mengutamakan penyiapan siswa untuk melanjutkan pendidikan kejenjang yang lebih tinggi dengan pengkhususan. Perwujudan pengkhususan tersebut berupa diselenggarakannya penjurusan yang dimulai di kelas X (sepuluh), yakni penjurusan

Matematika dan Ilmu Pengetahuan Alam, Ilmu-Ilmu Sosial, Bahasa, dan Agama.

Algoritma *k-means* adalah metode klasterisasi yang paling terkenal dan banyak digunakan diberbagai bidang karena sederhana, mudah diimplementasikan, memiliki perhitungan untuk mengklaster data yang besar, mampu menangani data banyak, dan kompleksitas waktunya linier  $O(nKT)$  dengan jumlah dokumen, K adalah jumlah kluster dan T adalah jumlah iterasi. Algoritma *k-means*

adalah metode pengklasteran secara data yang akan dipilih kedalam kelompok yang berbeda. Selama ini penjurusan siswa di MAN 1 JEMBER masih menggunakan perhitungan secara manual menggunakan microsoft excel untuk mencatat setiap nilai siswa baru.

## 1.2 Rumusan Masalah

Rumusan masalah dalam penelitian ini adalah:

- 1) Bagaimana menentukan kriteria penjurusan siswa baru di MAN 1 Jember?
- 2) Apakah metode *algoritma k-means clustering* dapat diterapkan dalam studi kasus penentuan penjurusan siswa MAN 1 Jember?
- 3) Apakah ada pengaruh perubahan centroid terhadap jumlah iterasi?

## 1.3 Batasan Masalah

Batasan Masalah dari penelitian ini adalah:

- 1) *Algoritma k-means* hanya diimplementasikan pada penjurusan siswa kelas X di MAN 1 JEMBER.
- 2) Data yang diambil hanya berdasarkan pada data siswa tahun 2014 kelas X dengan jumlah siswa 374 yang terbagi menjadi 11 kelas yaitu 5 Kelas MIPA, 4 Kelas IIS, 1 kelas Bahasa, dan 1 kelas Agama.

## 1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah:

- 1) Mengelompokkan siswa MAN 1 JEMBER dalam hal penjurusan siswa dengan menggunakan metode clustering *algoritma k-means*.
- 2) Menentukan pembagian jurusan menggunakan metode *algoritma k-means clustering* yang sesuai dengan kemampuan dari masing-masing siswa .

## 1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah:

- 1) Dapat membantu siswa dalam menentukan kriteria penjurusan.
- 2) Sistem dapat dijadikan sarana dalam penentuan jurusan sehingga menghindari ketidaktepatan dalam penjurusan berdasarkan analisis *algoritma k-means*.

## BAB II TINJAUAN PUSTAKA

### 2.1 Data Mining

*Data Mining* adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual.

#### 2.1.1 Tahapan Dalam Proses Data Mining

Dalam proses *data mining* ada beberapa tahapan, yaitu seleksi data, *pre-processing / cleaning, transformasi, data Mining, Interpretasi/Evaluasi*, Tahap Akhir

#### 2.1.2 Metode-Metode dalam Data Mining

*Data mining* dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu Klasifikasi, *Clustering, Association, Regression, Forecasting, Sequence Analysis, Deviation Analysis*.

### 2.2 Clustering

*Clustering* adalah proses mengelompokkan objek berdasarkan informasi yang diperoleh dari data yang menjelaskan hubungan antar objek dengan prinsip untuk memaksimalkan kesamaan antar anggota satu kelas dan meminimumkan kesamaan antar kelas/*cluster*.

#### 2.3 K-Means Clustering

*K-Means clustering* merupakan salah satu metode data *clustering* non-hirarki yang mengelompokkan data dalam bentuk satu atau lebih *cluster*/kelompok.

#### 2.4 Euclidean Distance

*Euclidean Distance* adalah metrika yang paling sering digunakan untuk menghitung kesamaan 2 vektor.

#### 2.5 Visual Basic.NET

*Microsoft Visual Basic .NET* adalah sebuah alat untuk mengembangkan dan membangun aplikasi yang bergerak di atas sistem *.NET Framework*, dengan menggunakan bahasa *Basic*.

#### 2.6 MySQL

*MySQL* adalah sebuah perangkat lunak system manajemen basis data SQL (*database management system*) atau DBMS yang *multithread* dan *multiuser*.

#### 2.7 Penjurusan siswa MAN 1 JEMBER

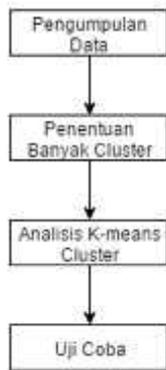
Penjurusan kelas X di MAN 1 Jember yang berjumlah 374 siswa, terdiri dari 5 Kelas MIPA, 4

Kelas IIS, 1 kelas Bahasa, dan 1 kelas Agama. Penjurusan MAN 1 JEMBER dimulai dari kelas X

### BAB III METODOLOGI PENELITIAN

#### 3.1 Desain Sistem

Desain sistem diperlukan untuk mempermudah proses dalam merancang suatu sistem.



Gambar 3.1 Blok Diagram Penelitian

##### 3.1.1 Teknik Pengumpulan Data

Menggunakan Metode Observasi, Metode Studi Kepustakaan, Metode Wawancara, Metode Dokumentasi

##### 3.1.2 Analisa Masalah

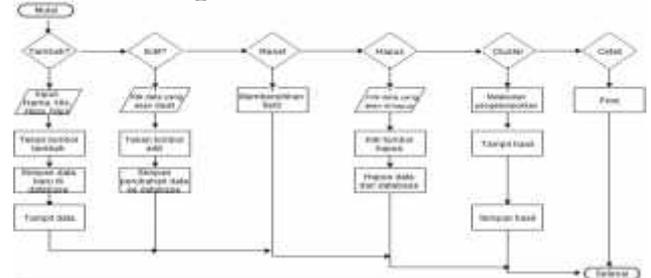
guru dalam membuat kelompok jurusan siswa yang masih dilakukan secara manual, hal ini disayangkan karena akan menyita banyak waktu guru untuk mengelompokkan siswanya, sehingga menjadikan faktor penghambat tugas-tugas yang seharusnya dapat terselesaikan dengan proses yang cepat, efektif, dan efisien. Untuk itu, sistem program yang akan dibuat ini akan berusaha membantu guru dalam memberikan saran penjurusan siswa sesuai dengan kemampuan siswanya dan memprediksi jumlah kelompok jurusan.

##### 3.1.3 Analisa Sistem

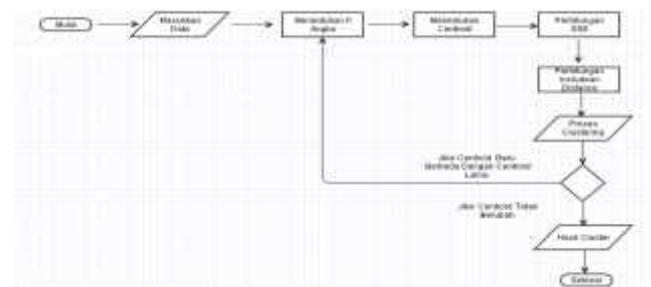
Dari hasil ujian tersebut dikelompokkan ke dalam 4 jurusan yaitu jurusan Matematika dan Ilmu Pengeahuan Alam (MIPA), Ilmu-Ilmu Sosial (IIS), Bahasa, dan Agama. Yang nantinya akan dijumlah melalui media *excel* sebelumnya. Setiap kelas penjurusan akan ditentukan pusat *cluster/centroid* yang berbeda untuk tiap-tiap kelompok yang nantinya akan didapatkan sebuah

keputusan penjurusan bagi tiap siswa melalui *centroid* setiap iterasi.

### 3.2 Perancangan Sistem



Gambar 3.2 Flowchart Antarmuka Perangkat Lunak



Gambar 3.3 Flowchart algoritma k-means

### BAB IV HASIL DAN PEMBAHASAN

Dalam bab ini menjelaskan tentang implementasi dari penjurusan siswa di MAN 1 Jember menggunakan *K-Means*.

#### 4.1 Pembahasan Hasil

##### 4.1.1 Hasil Program

Menu Utama merupakan tampilan saat aplikasi dijalankan, serta seluruh tampilan menu yang ada pada program, tampilan sebagai berikut:



Gambar 4.1.1 Tampilan Menu Aplikasi Penjurusan Siswa



Gambar 4.1.2 Tampilan Input Siswa

Tombol “Simpan” untuk menambahkan data pada aplikasi, jika tombol “Simpan” di klik dan data baru yang akan di Simpankan tersebut belum di entrikan kedalam form pengisian maka akan muncul peringatan pesan atau peringatan seperti gambar di bawah ini:



Gambar 4.1.3 Pesan Peringatan Simpan Data

Dari inputan data dapat diketahui jumlah data terdapat 374 siswa. Setelah di klik tombol “Simpan” maka data secara otomatis akan masuk pada database, berikut tampilannya:



Gambar 4.1.4 Database K-Means

### 4.1.2 Menu Perhitungan K-Means

Menu Perhitungan K-Means merupakan tampilan setelah aplikasi dijalankan yaitu input siswa, serta seluruh tampilan menu yang ada pada program.



Gambar 4.1.5 Perhitungan K-Means

Menu ini berisi tentang perhitungan K-Means setelah menu input siswa diisi akan muncul perhitungan K - Means. Nantinya perhitungan ini akan menghitung sampai mana proses iterasi K-Means ini. Jika iterasi dan posisi sudah ada yang sama maka proses perhutungan K-Means berhenti. Dalam menu perhitungan K-Means ini juga terdapat Detail Iterasi untuk melihat perubahan posisi setiap iterasi.



Gambar 4.1.6 Tampilan Menu Detail Perhitungan Tiap Iterasi

### 4.1.3 Menu Hasil Perhitungan K-Means



Gambar 4.1.6 Tampilan Menu Hasil MIPA



Gambar 4.1.7 Tampilan Menu Hasil IIS



Gambar 4.1.8 Tampilan Menu Hasil BAHASA



Gambar 4.1.9 Tampilan Menu Hasil AGAMA



Gambar 4.1.10 Tampilan Menu Hasil Semua

NAMA	MIPA	IIS	Bahasa	Agama
AKHMAD MA'LUFIL WARO	98	82	83	83
ALVIAN IQBAL FAHMI	76	75	85	87
FARIEZUL HAZNAM	87	81	80	89
DEPHI IRFANI	84	78	88	82
MUHAMMAD FIRRI AL AEFKONI	85	88	82	80
SHENI IKHSAN	82	78	71	82
AFNA YANUI ALFI LAYALIN	75	78	81	79
ALVIN RIDHANI NGTYAS	88	88	88	88
ANINDI ZAHROTIL FAUZHAH	82	82	87	82
ANGGUN NUI CAHYANI	88	72	87	88

## 4.2 Pengujian Dan Evaluasi

Pada tahap ini, rancangan perangkat lunak diterapkan sebagai sebuah program atau unit program.

### 4.2.1 Penentuan Centroid

*Centroid* adalah rata-rata (*mean*) dari setiap titik anggota *cluster*. Untuk Menentukan *K* angka yaitu dengan menentukan berapa jumlah kelas yang akan digunakan dalam hal ini antara lain MIPA, IIS, Bahasa, dan Agama. Kemudian menentukan *centroid* yang dipakai dengan rumus perhitungan SSE (*Sum Square Error*) yaitu perhitungan untuk menentukan *centroid* awal (SSEmax) yang berdasarkan nilai tertinggi. Lalu SSE dari setiap cluster dapat menjadi ukuran performansi baik/tidaknya *cluster* tersebut. Dari dua jumlah cluster yang telah dicoba, dapat diketahui bahwa SSE dari jumlah cluster 1 (MIPA) lebih kecil dibandingkan dengan jumlah cluster 4 (Agama). Ini menunjukkan entitas yang masuk pada jumlah cluster 1 memiliki kesamaan yang lebih seragam dari pada entitas yang ada pada jumlah cluster 4. Maka diketahui *centroid* yang dipakai penjurusan dari semua siswa MIPA sampai Agama menggunakan nilai tertinggi.

### 4.2.2 Perhitungan K-Means

Dalam pembahasan ini akan dijelaskan perhitungan manual Algoritma *K-Means*. Bagaimana alur dan cara perhitungan sistem terhadap data siswa MAN 1 Jember yang akan meneliti 374 siswa sesuai dengan ketentuan di sekolah tersebut.

*Centroid* awal di tentukan dengan nilai yang paling tinggi di masing-masing kelompok. Seperti nilai di atas :

$$C \text{ Nilai MIPA} = \{98, 0, 0, 0\}$$

$$C \text{ Nilai IIS} = \{0, 90, 0, 0\}$$

$$C \text{ Nilai Bahasa} = \{0, 0, 98, 0\}$$

$$C \text{ Nilai Agama} = \{0, 0, 0, 93\}$$

Sebagai contoh kita ambil data nilai dari siswa yang berama:

AKHMAD MA'LUFIL WARO 98 82 83 83

menghitung jarak setiap data terhadap *centroid* perhitungannya seperti dibawah ini :

#### MIPA

$$\begin{aligned} &= \sqrt{(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2 + (w - w_1)^2} \\ &= \sqrt{(98 - 98)^2 + (82 - 0)^2 + (83 - 0)^2 + (83 - 0)^2} \\ &= \sqrt{(0)^2 + (82)^2 + (83)^2 + (83)^2} \\ &= \sqrt{0 + 6724 + 6889 + 6889} \\ &= \sqrt{20502} \\ &= 143,1852 \end{aligned}$$

#### IIS

$$\begin{aligned} &= \sqrt{(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2 + (w - w_1)^2} \\ &= \sqrt{(98 - 0)^2 + (82 - 90)^2 + (83 - 0)^2 + (83 - 0)^2} \\ &= \sqrt{(98)^2 + (-8)^2 + (83)^2 + (83)^2} \\ &= \sqrt{9604 + 64 + 6889 + 6889} \\ &= \sqrt{23446} \\ &= 153,1209 \end{aligned}$$

#### Bahasa

$$\begin{aligned} &= \sqrt{(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2 + (w - w_1)^2} \\ &= \sqrt{(98 - 0)^2 + (82 - 0)^2 + (83 - 98)^2 + (83 - 0)^2} \\ &= \sqrt{(98)^2 + (82)^2 + (-15)^2 + (83)^2} \\ &= \sqrt{9604 + 6724 + 225 + 6889} \\ &= \sqrt{23442} \\ &= 153,1078 \end{aligned}$$

#### Agama

$$\begin{aligned} &= \sqrt{(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2 + (w - w_1)^2} \\ &= \sqrt{(98 - 0)^2 + (82 - 0)^2 + (83 - 0)^2 + (83 - 93)^2} \\ &= \sqrt{(98)^2 + (82)^2 + (83)^2 + (-10)^2} \\ &= \sqrt{9604 + 6724 + 6889 + 100} \\ &= \sqrt{23317} \\ &= 152,6991 \end{aligned}$$

Jarak terhadap *centroid* :

$$C_{mipa} = 143,1852$$

$$C_{iis} = 153,1209$$

Cbahasa= 153,1078  
Cagama= 152,6991

Untuk proses selanjutnya adalah proses pengelompokan atau *clustering* algoritma *k-means*. dengan hasil di atas jarak terdekat dengan *centroid* adalah nilai jarak yang terkecil yang akan di pilih untuk di kelompokkan, data di atas menunjukkan bahwa nilai terkecil terdapat pada nilai "MIPA" dan "AGAMA", maka data tersebut tercluster ke kelas "MIPA" dan "AGAMA".

### 4.3 Uji Centroid

Dalam uji *centroid* ini akan menghitung *centroid* 1 sampai *centroid* 10 dengan melakukan data 100 siswa dan dibagi menjadi kelompok. Kemudian *centroid* 1 dilakukan dengan data 1 sampai 10 siswa, *centroid* 2 data 11 sampai 20 siswa. Untuk posisi berubah atau tidak dalam pengerjaan tugas akhir ini, saya memilih *centroid* 1 dengan nilai iterasi 15.

## BAB V PENUTUP

### 5.1 Kesimpulan

Dari hasil uji coba penelitian kali ini penulis mencoba memaparkan beberapa kesimpulan antara lain:

1. Untuk hasil manual dan otomatis didapat bahwa dari 100 data, ada 69 data yang memiliki kesamaan, dan 31 data yang tidak memiliki kesamaan. Sehingga prosentase tingkat erornya adalah 44,9%.
2. Penentuan penjurusan siswa MAN 1 Jember berdasarkan jurusannya terbagi menjadi 4 *cluster*, yaitu *kluster* K1 (MIPA), *kluster* K2 (IIS), *kluster* K3 (Bahasa), dan *kluster* K4 (Agama) yang telah dilakukan penghitungan dengan menggunakan algoritma *K-Mens Clustering*. Sehingga dapat diketahui pada *cluster* ke berapa penempatan siswa penjurusan berdasarkan kelas penjurusannya di MAN 1 Jember.
3. Perubahan *centroid* memiliki pengaruh terhadap iterasi, karena setiap kelas penjurusan akan ditentukan pusat *cluster/centroid* yang berbeda untuk tiap-tiap kelompok yang nantinya akan didapatkan sebuah keputusan penjurusan bagi tiap siswa melalui *centroid* setiap iterasi.

### 5.2 Saran

Adapun saran agar dapat meningkatkan aplikasi ini dimasa mendatang adalah sebagai berikut :

1. Menambahkan variabel lain misal seperti ekstrakurikuler dan hasil test psikologis dalam menentukan penjurusan siswa.
2. Menambahkan tampilan aplikasi yang lebih menarik dan interaktif.
3. Sistem dirancang dengan berbasis web agar dapat di akses oleh siswa dan orang tua di dalam ataupun diluar sekolah.
4. Menambahkan menu yang lebih banyak seperti import data dari excel atau pdf.
5. Menggunakan metode lain untuk menghasilkan tingkat akurasi yang lebih besar.

### DAFTAR PUSTAKA

Agusta, Yudi. 2007. "*K-Means – Penerapan, Permasalahan dan Metode Terkait*". STMIK STIKOM BALI..

Bachtiar.2010. "*Rancang Bangun Sistem Pendukung Keputusan (SPK) Penjurusan Program Studi*". Fakultas Sains dan Teknologi, Universitas Islam Negeri Syarif Hidayatullah. Jakarta

Eric Budiman, Isye Arieshanti dan Rully Soelaiman. 2013. "*Implementasi KD-Tree-K-Means Clustering untuk Klasterisasi Dokumen*". Fakultas Teknologi Informasi Institut Teknologi Sepuluh Nopember (ITS).

Gautama, Mohammad Glesung. 2010. "*Penentuan Jurusan Di SMAN 8 Surakarta Dengan Fuzzy Inference System (FIS) Mamdani*". Fakultas Matematik dan Ilmu Pengetahuan Alam. Universitas Sebelas Maret. Surakarta.

Kusrini, Emha Taufik, Luthfi. 2009. "*Algoritma Data Mining*". Yogyakarta: Andi

Setiawan, Deni. 2015. "*Perancangan Aplikasi K-Means Sebagai Penentu Konsentrasi Bagi Mahasiswa Informatika UMS*". Fakultas komunikasi dan Informatika, Universitas Muhammadiyah. Surakarta.

Tedy dan Sri Kusumadewi. 2008. "*Aplikasi Untuk Mengelompokkan Mahasiswa Berdasarkan Nilai Body Mass Index (Bmi) & Ukuran Kerangka*". Jurusan Teknik Informatika FTI, UII. Yogyakarta.

Teguh Hariyadi, Suprayogi. 2012. "*Penerapan Algoritma K-Means Untuk Pengelompokan Data Nilai Siswa*". Fakultas Ilmu Komputer, Universitas Dian Nuswantoro. Semarang

Teknomo, Kardi. 2006. "*K-means Clustering Tutorial.*" *Medicine*. 100.4 (2006): 3.

Turban, Efraim, Kelly Rainer, Richard E Potter. 2005. "*Pengantar Teknologi Informasi*". Jakarta:Salemba.

Wulandari, Retno. 2012. "*Pengenalan rumput laut menggunakan Euclidean distance berbasis Ekstra fitur*". Universitas Nusantara PGRI. Kediri.

Yakub. 2012."Pengantar Sistem Informasi". Yogyakarta: Graha Ilmu.



