

OPTIMASI PENENTUAN CENTROID PADA ALGORITMA *K-MEANS* UNTUK PEMILIHAN JURUSAN BERBASIS GENETIKA ALGORITMA

ACHMAD BAIJURI¹, AHMAD LUTFI²

Sistem Informasi, Fakultas Sains dan Teknologi, Universitas Ibrahimy,
Banyuputih (68373), Situbondo, Indonesia¹. e-mail: bayubai@gmail.com

Sistem Informasi, Fakultas Sains dan Teknologi, Universitas Ibrahimy,

Banyuputih (68373), Situbondo, Indonesia². e-mail:

ahmadlutfi.14@gmail.com

ABSTRAK

Muhasabah bukanlah hal yang sangat mudah dilakukan tanpa niat yang kuat. Apalagi jika dikhususkan pada Muhasabah kemampuan dan Minat yang dimiliki diri sendiri, termasuk juga Penentuan Jurusan terhadap beberapa siswa, bukanlah hal yang mudah. Banyak siswa yang belum mengenal kemampuan dan minat yang dimilikinya. Civitas Smk Khamas juga kebingungan dalam memeberikan sebuah pelajaran ekstra terhadap siswa. Dari masalah tersebut sangatlah dibutuhkan metode yang bisa memetakan minat siswa sehingga siswa tidak salah pilih jurusan. Algoritma K-Means Algoritma yang bisa memetakan / mengelompokan siswa yang mempunyai kemiripan sama. Algoritma K-Means sangat sederhana untuk diimplementasikan serta mudah dijalankan, serta relatif cepat dalam menjalankan. Namun ada masalah dalam penentuan centroid sehingga dibutuhkan optimasi dalam hal tersebut.

Kata kunci : *K-Means*; penjurusan siswa; *Genetika Algoritma*; Klaster

I. PENDAHULUAN

Muhasabah bukanlah hal yang sangat mudah dilakukan tanpa niat yang kuat. Apalagi dikhususkan terhadap Muhasabah kemampuan dan Minat yang dimiliki sendiri. Muhasabah sangatlah penting bagi setiap insan Manusia. Dengan Muhasabah, insan manusia dapat tahu akan kelebihan dan kekurangan dirinya. Dengan begitu, insan manusia dapat lebih mudah dalam memperbaiki kekurangan serta pengoptimalan kelebihan diri.

Penentuan jurusan terhadap siswa bukanlah hal yang mudah. Banyak siswa yang belum mengenal minatnya serta kemampuan yang dimiliki dirinya. Hal tersebut membuat siswa cenderung memilih dan menjalani jurusan yang kurang sesuai kemampuan dan minatnya.

SMK Khamas adalah salah satu Sekolah yang memiliki 2 Jurusan, yakni multimedia dan Teknik Komputer dan Jaringan di lingkungan Pondok Pesantren Salafiyah Syafi'iyah Nurul Arsyad Asembagus. Sekolah ini memiliki beberapa kelas dalam satu tingkat. Tentunya hal tersebut memerlukan banyak tenaga pengurus/ karyawan dalam melayani kegiatan belajar mengajar siswa.

Kegiatan aktifitas KBM kadang sering terganggu dengan adanya kekurangan tenaga Kependidikan, salah-satunya adalah aktifitas penentuan Tes Jurusan. Karena dalam kegiatan tersebut biasanya juga diadakan aktifitas Praktek Kerja Lapangan siswa kelas XI.

Selain dari masalah tersebut, penentuan Jurusan, Civitas Smk Khamas juga kebingungan dalam memeberikan sebuah Ekstrakurikuler guna meningkatkan skil terhadap siswa. Tentunya hal tersebut dikarenakan kurang begitu paham terhadap kelompok siswa yang kemampuannya kurang baik di bidangnya.

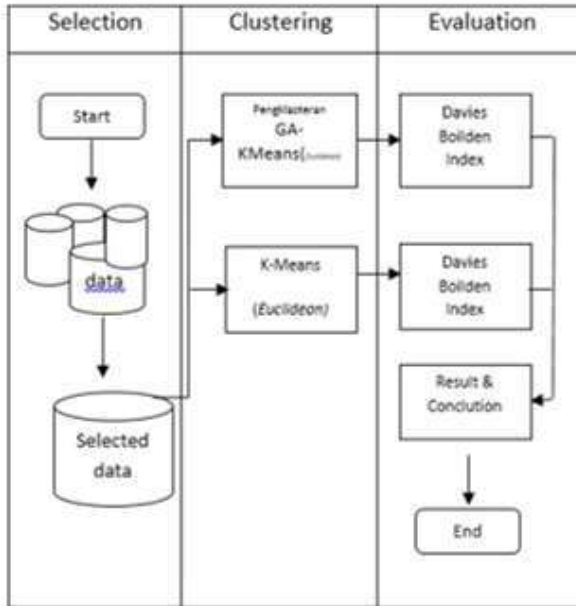
Dari permasalahan yang dijelaskan di atas, kiranya peneliti beranggapan perlu adanya SPK (sistem Pengambil Keputusan) tentang penentuan jurusan terhadap siswa baru. Peneliti berusaha menyelesaikan masalah ini dengan sistem komputer untuk memproses cepat dan memudahkan dalam penentuan jurusan.

K-Means Merupakan salah satu metode pengelompokan data nonhierarki (sekatan) yang berusaha mempartisi data yang ada kedalam bentuk dua tau lebih kelompok. Metode ini mempartisi data ke dalam kelompok sehingga data berkarakteristik sama dimasukkan dalam satu kelompok yang sama dan data yang berkarakteristik berbeda dikelompokkan ke dalam kelompok yang lain. Adapun tujuan pengelompokan data ini adalah untuk meminimalkan fungsi objektif yang diset dalam proses pengelompokan, yang pada umumnya berusaha meminimalkan variasi di dalam suatu kelompok dan memaksimalkan variasi antar kelompok. Kelebihan dari metode K-Means adalah mudah di implementasikan, dapat mengelompokkan objek besar dan outlietnya sangat cepat.

Rumuskan masalah yaitu mengelompokan siswa yang berkarakteristik sama di SMK Khamas menggunakan K-Means dengan mengoptimasi penentuan centroidnya sehingga terklaster dan bermanfaat dalam penentuan jurusan.

II. METODE PENELITIAN

Tahapan penelitian yang dilakukan adalah pengumpulan data awal, Seleksi, Klustering, evaluasi dan Validasi, dan berakhir dengan hasil dari penelitian. Alur tahapan metode yang diusulkan pada penelitian ini dijelaskan sebagaimana pada Gambar 1 berikut ini:



Gambar 1. Alur Metode yang Diusulkan

1. *Selection*

Dataset awal diambil dari data mahasiswa pada waktu Tespenmaru di Smk Khamas Asembagus Situbondo. Dari beberapa data diperbaiki dan dipilih (*selection*) dan digabung data-data yang diperlukan untuk penelitian ini (*selected data*).

2. *Clustering*

Pengklasteran dataset mahasiswa digunakan dengan perbandingan dua metode, tahap pertama dilakukan tanpa menggunakan optimasi (*optimasi GA*), tahap kedua dilakukan dengan *GA K-Means*.

3. Evaluasi dan Validasi

Untuk mengetahui validitas hasil eksperimen digunakan uji validitas *Davies-Bouldin Index*. Proses uji akan menghasilkan nilai skema *Clustering*. Nilai yang paling minimal adalah nilai *Davies Bouldin-Index* yang paling optimal.

III.HASIL PEMBAHASAN

1. Selection

Data mahasiswa yang ada dikonfersi menjadi angka terlebih dahulu, kemudian dipilih sesuai atribut yang akan digunakan. Data yang sudah jadi dilihat pada Tabel 1

Tabel 1. Data Siswa

NAMA MAHASISWA	Jaringan	Mtk	Simdik	PMD	Desain Grafis	Audio Video	Siskom	Sistem Operasi
Responden 1	74	70	70	70	79	79	79	79
Responden 2	70	70	70	70	79	74	79	79
Responden 3	74	70	70	58	74	74	79	79
Responden 4	70	70	70	58	74	74	79	79
Responden 5	70	70	58	58	74	79	79	79
Responden 6	70	70	70	58	74	70	79	79
Responden 7	74	70	70	58	74	79	79	70
Responden 8	74	70	79	79	74	70	79	79
Responden 9	70	70	70	70	74	70	79	79
Responden 10	74	74	74	79	74	79	79	79
Responden 11	70	74	70	58	74	74	79	79
Responden 12	74	58	79	79	79	74	79	79
Responden 13	74	79	70	79	70	74	74	79
Responden 14	70	70	70	58	70	74	79	79
Responden 15	70	70	70	79	79	74	79	79
Responden 16	70	70	70	58	74	74	79	79
Responden 17	70	70	70	79	79	79	79	79
Responden 18	74	70	70	58	79	70	79	79
Responden 19	58	70	70	58	79	70	79	79
Responden 20	70	70	70	70	79	79	74	79
Responden 21	70	70	70	58	70	79	74	79
Responden 22	74	70	79	58	74	79	79	70

Responden 23	62	70	79	70	70	79	79	70
Responden 24	70	79	70	70	70	79	79	70
Responden 25	58	70	70	70	74	79	79	70
Responden 26	58	70	70	58	74	74	62	70
Responden 27	62	70	58	58	74	74	79	70
Responden 28	62	70	58	58	44	79	74	70
Responden 29	62	70	58	58	54	79	79	70
Responden 30	70	70	58	58	79	79	79	70
Responden 31	62	70	79	58	70	79	79	70
Responden 32	70	70	58	58	74	62	79	70
Responden 33	74	70	70	58	70	79	79	70
Responden 34	74	70	58	58	70	74	79	70
Responden 35	70	70	70	58	74	79	79	58
Responden 36	70	70	70	58	79	79	79	70
Responden 37	62	70	58	58	70	79	79	70

2. Clustering K-Means dan Evaluasi DBI

Implementasi Clustering dilakukan pada data yang di table 4.1. data tersebut diklaster menggunakan *K-Means* kemudian Optimasi *GA K-Means* Penentuan jumlah cluster untuk pengolahan pertama yakni $K=3$. Implementasi ke 2 dilakukan dengan Optimasi *GA* dan $K=3$ juga. Hasil dari proses perhitungan menggunakan *K-Means* yakni Anggota klaster 1 sebanyak enam belas Mahasiswa, klaster 2 sebanyak 11(sebelas) mahasiswa, klaster 3 sebanyak 10 (sepuluh) Mahasiswa. sedangkan dengan menggunakan Algoritma *GA K-Means* yakni Anggota klaster 20 sebanyak 10 Mahasiswa, klaster 2 sebanyak 11 mahasiswa, klaster 3 sebanyak 7 Mahasiswa.

Hasil perhitungan nilai DBI masing-masing dataset adalah sebagai berikut:

Tabel .2. Nilai DBI yang dihasilkan pada *K-Means*

R	C1	C2	C3	R MAX	DBI
C1	0	16,0529	25,228	25,228	22,1696
C2	16,0529	0	13,8218	16,0529	
C3	25,228	13,8218	0	25,228	

Tabel 3. Nilai DBI yang dihasilkan pada *GA K-Means*

R	C1	C2	C3	R MAX	DBI
C1	0	20,6062	22,0714	22,0714	21,583
C2	20,6062	0	7,45242	20,6062	
C3	22,0714	7,45242	0	22,0714	

Dari hasil diatas menunjukkan DBI Optimasi GA menunjukan 21,5823, dan hal itu lebih kecil dari DBI tanpa menggunakan Optimasi. Dengan demikian sesuai dengan pendapat ahli yang disebutkan di landasan teori bahwa Semakin kecil nilai *DB Index* akan menunjukan skema *cluster* yang paling optimal. Maka dengan begitu Genetika Algoritma berjalan secara optimal.

IV.KESIMPULAN

Dari hasil penelitian ini dapat disimpulkan sebagai berikut:

Telah dilakukan pengklasteran siswa dengan menggunakan algoritma K-means. Pengklasteran menghasilkan 3 klaster. Anggota klaster 1 sebanyak 16 siswa, klaster 2 sebanyak 11 siswa, klaster 3 sebanyak 10 siswa dengan menggunakan Algoritma *K-Means*. Sehingga dalam 3 klaster ini nantinya akan dijuruskan sesuai dengan kelompok yang sekiranya cocok terhadap jurusan

Pengklasteran ini telah mencapai nilai optimal pada iterasi keempat untuk Algoritma *K-Means*. Sedangkan ketika dioptimasi iterasi mencapai nilai Optimal pada iterasi 2.

Pengelompokan lebih optimal ketika menggunakan *Genetika Algoritma* yaitu dengan nilai validasi DBI 21,5823 sedangkan yang menggunakan K-means saja menunjukan Nilai DBI 22,1696.

DAFTAR PUSTAKA

- E. Prasetyo, *Data Mining Konsep dan Aplikasi menggunakan Matlab*, Yogyakarta: Andi Offset, 2012.
- R. Herawati, "Rekomendasi Penjurusan di SMU YSKI dengan Algoritma K-Means," *Proxies*, pp. 3-4, 2012.
- O. dan O. Eyelade, "Application of K-Means Clustering algorithm for prediction of Students Academic Performance," *IJCSIS*, vol. 7, pp. 2-3, 2010.
- S. Ray dan R. H. Turi, "Determination of number of clusters in k-means clustering and application in colour image segmentation," p. 3, 1999.
- B. Al-Shboul dan S. H. Myaeng, "Initializing K-Means using Genetic Algorithms," *World Academy of Science, Engineering and Technology*, vol. 3, p. 6, 2009.
- R. D. Ramdhani, "Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Universitas Dian Nuswantoro," p. 2, 2014.
- Widiarina dan R. S. Wahono, "Algoritma Cluster Dinamik untuk Optimasi Cluster pada Algoritma K-Means dalam Pemetaan Nasabah Potensial," *Journal of Intelligent Systems*, 2015.
- F. A. Hermawati, *Data Mining*, Yogyakarta: Andi Offset, 2013.
- H. W. Nisa dan A. S. Rakhman, "Fuzzy Inference dengan metode Tsukamoto sebagai pemberi saran Pemilihan Konsentrasi," *Snati*, p. 3, 2012.
- Y. Nugroho, "Klasifikasi dan Klastering Penjurusan Siswa SMA Negeri 3 Boyolali," *Khazanah Informatika*, vol. 1, 2016.
- T. Rismawan dan S. K. Dewi, "Pengelompokan Mahasiswa berdasarkan BMi dan ukuran kerangka," *Journal its*, pp. 4-5, 2010.
- F. K. Wardani, "Penerapan Metode GA-Kmeans untuk pengelompokan pengguna pada Bapersip Provinsi Jawa Timur," *Journal Teknik ITS*, vol. 1, pp. 1-2, 2012.

