

Clustering Data Nilai Siswa SMA Untuk Penentuan Jurusan Menggunakan Algoritma Fuzzy C-Means

Fajar Akbar¹⁾

¹⁾ Program Pascasarjana Magister Ilmu Komputer STMIK Nusa Mandiri
Jl. Salemba Raya No. 5 Jakarta Pusat
Fajar.fkb@bsi.ac.id

ABSTRACT

Determination of the majors a process that occurs at the time the class- X will go up to class- IX, this is because the addition proses is carried out must be based in the interests and talent of each student's. In addition, this process is done by Bp/Bk, witch this assessment process based on value evaluation of each subject's flagship from each Department, to major in Science related subject (biology, physics, chemistry, mathematics), to social scienece Departement (economics, sociology, geography, history), and for the Departement of IPB (Indonesian Language, English, Germany Language). In this reaseach FCM Clustering can assist and improve value for Determination of the majors process 72%.

KeyWords:

Fuzzy C-Means, Department, Clustering

ABSTRAKSI

Penentuan jurusan merupakan proses yang terjadi pada saat siswa kelas-X akan naik ke kelas- IX, ini dikarenakan proses penjurusan ini dilakukan harus berdasarkan minat dan bakat dari setiap siswa yang ada. Proses penjurusan ini dilakukan oleh bagian Bp/Bk, yang mana proses penilaian ini berdasarkan nilai evaluasi dari setiap mata pelajaran unggulan dari tiap-tiap jurusan, untuk jurusan mata pelajaran yang terkait IPA (Biologi, Fisika, Kimia, Matematika), untuk jurusan IPS (Ekonomi, Sosiologi, Geografi, Sejarah), dan untuk jurusan IPB (Bahasa Indonesia, Bahasa Inggris dan Bahasa Jerman). Didalam penellitian ini clustering dapat membantu dan meningkatkan ketepatan pemilihan jurusan atau pemilihan kosentrasi sebesar 72 % .

Kata Kunci:

Fuzzy C-Means, Jurusan, klustering

PENDAHULUAN

Sesuai dengan Permendikbud 81A tahun 2013 tentang implementasi kurikulum 2013 menjelaskan bahwa pendidikan merupakan proses yang sistematis untuk meningkatkan martabat manusia secara holistik yang memungkinkan potensi diri berkembang secara optimal. Peminatan atau yang lebih disebut penjurusan bertujuan untuk lebih memfokuskan atau mengarahkan materi pembelajaran para siswa-siswi sekolah menengah agar sesuai dengan minat dan bakat yang dimiliki oleh para siswa-siswi sekolah. Peminatan atau penjurusan ini juga membantu memberi gambaran akan bidang yang nantinya akan ditekuni oleh para siswa-siswi setelah lulus sekolah menengah nanti. Jadi para siswa-siswi yang telah lulus tidak kesulitan menentukan bidang yang akan ditekuni ketika kuliah atau ketika bekerja nanti.

Dengan metode *clustering* data yang ada akan dikelompokkan menjadi beberapa kluster yang masing-masing diwakili oleh pusat kluster. Data yang diklasifikasikan berdasarkan hasil akhir dari evaluasi nilai yang diperoleh siswa dan siswi berdasarkan setiap mata pelajaran yang mewakili setiap jurusan. Sedangkan “Algoritma Genetika banyak digunakan untuk masalah optimalisasi yang sulit dilakukan oleh metode konvensional” (Arochman, 1-9). “Metode ini dalam proses penyelesaiannya memformulasikan terminologi genetika yang diterapkan kedalam permasalahan sehari – hari”. (Wijaya, Bima Sukma dan Windarto, 189-197).

Tujuan dari penelitian ini diharapkan dapat membantu menyelesaikan permasalahan pemilihan jurusan agar tidak ada

kealahan dalam menentukan penjurusan bagi para siswa dan memperoleh hasil pengelompokan jurusan yang sesuai dan tepat dengan minat dan potensi setiap siswa-siswi agar bisa mengembangkan diri sebelum terjun langsung dalam bermasyarakat.

BAHAN DAN METODE

A. Pendekatan penjurusan

Sesuai dengan Pasal 12 ayat 2 Undang – Undang No 20 Tahun 2003 tentang Sistem Pendidikan yang menyatakan bahwa Setiap Peserta didik berhak mendapatkan pelayanan pendidikan sesuai dengan bakat, minat dan kemampuannya. Pasal ini menjelaskan bahwa setiap siswa sekolah, khususnya sekolah menengah atas ataupun sederajat mempunyai hak dalam menentukan pilihan pembelajaran sesuai dengan potensi, minat dan bakat yang dimiliki oleh para siswa-siswi. Pihak sekolah wajib menyediakan jurusan atau peminatan agar siswa-siswi dapat melanjutkan pengembangan potensi yang dimiliki oleh masing – masing individu.

Proses pemilihan peminatan pada dilakukan pada saat kenaikan kelas dari kelas X ke Kelas XI. Proses peminatan yang dilakukan berdasarkan nilai evaluasi dari setiap mata pelajaran peminatan, selain nilai evaluasi dari setiap mata pelajaran peminatan sebagai bahan pertimbangan juga dilihat dari minat dari siswa/i yang bersangkutan.

B. Konsep *Clustering* Pada *Data Mining*

Proses klustering pada data mining menurut Hermawati (2009a:16) “Proses mempartisi data-set menjadi beberapa sub-set atau kelompok sedemikian

rupa sehingga elemen-elemen dari suatu kelompok tertentu memiliki set properti yang *dishare* bersama”.

Analisa klastering Hermawati(2009b:123) merupakan analisa yang dilakukan pada sekumpulan objek sehingga objek – objek dalam satu kelompok sama dengan yang lain dan berbeda dengan objek dalam kelompok yang lainnya. Tujuan dari melakukan analisa klaster ialah meminimalkan jarak didalam klaster dan memaksimalkan jarak antara klaster. Tipe-tipe dari Clustering sebagai berikut:

1. *Well –Separated Cluster*
2. *Center Based Cluster*
3. *Contiguous Cluster (nearest neighbor atau Trasnsitive*
4. *Density Based*

C. Algoritma Fuzzy C-Means

Fuzzy C-means merupakan suatu teknik peng-klasteran data yang keberadaan setiap titik data dalam suatu klaster ditentukan oleh derajat keanggotaanya (Purnomo dan Kusumadewi:2010). Konsep dasar FCM yaitu menentukan pusat klaster, yang akan menandai lokasi rata-rata untuk setiap klaster. Dengan caramemperbaiki pusat klaster akan bergerak menuju lokasi yang tepat. Perulangan ini didasarkan pada minimasi fungsi obyektif yang menggambarkan jarak dari titik data yang diberikan kepusat klaster yang berbobot oleh derajat keanggotaan titik data tersebut.

Algoritma FCM didasarkan pada minimasi fungsi obyektif yang diformulasikan dalam persamaan:

$$J_m = \sum_{i=1}^N \sum_{j=1}^C (u_{ij})^m ||z_i - c_j||^2$$

Dimana:

U_{ij} = Level keanggotaan dari Z_i dalam cluster j .

Z_i = Nilai data ke-I dari d- dimensi data.

C_i = Nilai ke-j dari d-dimensi cluster center

m = Sembarang bilangan real lebih besar dari 1

Tahapan penyelesaian Algoritma FCM disusun dengan langkah sebagai berikut:

Langkah 1: Tentukan himpunan data Z . tentukan jumlah *cluster* yang diharapkan $1 < c < N$, nilsi pembobot $m > 1$, toleransi penghentian $\epsilon > 0$.

Langkah 2: Inisialisasi matriks partisi secara acak, $U(0) \in M_{fc}$

Ulangi untuk $l=3,4,5$

Langkah 3: Hitung *cluster center (means)*.

$$v_i^{(l)} = \frac{\sum_{k=1}^N \left(\mu_{ik}^{(l-1)} \right)^m z_k}{\sum_{k=1}^N \left(\mu_{ik}^{(l-1)} \right)^m} \dots 1 \leq i \leq c$$

Langkah 4: Hitung jarak.

Langkah 5: Perbaharui matriks partisi

Untuk $1 \leq k \leq N$

Jika $D_{ikA} > 0$ untuk semua $i=1,2,3,\dots,c$.

$$\mu_{ik}^{(t)} = \frac{1}{\sum_j^c \left(D_{ikA} \frac{D_{jkA}}{D_{ikA}} \right)^{\frac{2}{m-1}}}$$

Atau dengan kata lain:

$$\mu_{ik}^{(t)} = 0 \text{ jika } D_{ikA} > 0 \text{ dan } \mu_{ik}^{(t)} \in [0,1]$$

$$\text{dengan } \sum_{i=1}^c \mu_{ik}^{(t)} = 1$$

Ulangi sampai dengan $\|U(t) - U(t-1)\| < \epsilon$.

D. Metode Penelitian Eksperimental

Penelitian yang peneliti lakukan merupakan penelitian jenis eksperimental, dimana menurut F. Hair, Jr., Rolph, E. A., Romald, L.T. dan G.B William (2008) dalam Sangadji etta dan Sopian (2010). Penelitian eksperimental (*Experimental Research*) merupakan penelitian yang subjeknya diberi perlakuan (*treatment*) lalu diukur akibat perlakuan pada diri subjek. tujuan dari penelitian ekperimental untuk melihat pengaruh variable independen terhadap variable dependen.

Peneliti dalam penelitian ini membahas tentang bagaimana algoritma Fuzzy C-Means (FCM) mampu memberikan rekomendasi pengelompokan atau penentuan jurusan yang optimal bagi para siswa-siswi sekolah dan juga bagaimana

Algoritma Gentika mendistribusikan $D_{ikA}^2 = [(z_k v_i^{(t)})^T A (z_k - v_i^{(t)})]$, $1 \leq i \leq c, 1 \leq k \leq N$ siswa kedalam kelas berdasarkan kuota yang tersedia. Proses pengelompokan atau penentuan jurusan ini didapat berdasarkan evaluasi nilai setiap mata pelajaran jurusan dan berdasarkan hasil evaluasi nilai permata pelajaran jurusan.

E. Ekperimen

Data populasi yang peneliti ambil meupakan data populasi sebaran nilai hasil evaluasi siswa-siswi kelas- X semester genap pada tahun 2013. Jumlah populasi didalam data hasil evaluasi nilai siswa-siswi berjumlah 280 Orang, dari jumlah tersebut akan ambil sampel sejumlah 155 Orang. Jumlah ini diperoleh berdasarkan table yang dikembangkan dari *Isaac* dan *Michael* dengan tingkat kesalahan, 5%.

HASIL DAN PEMBAHASAN

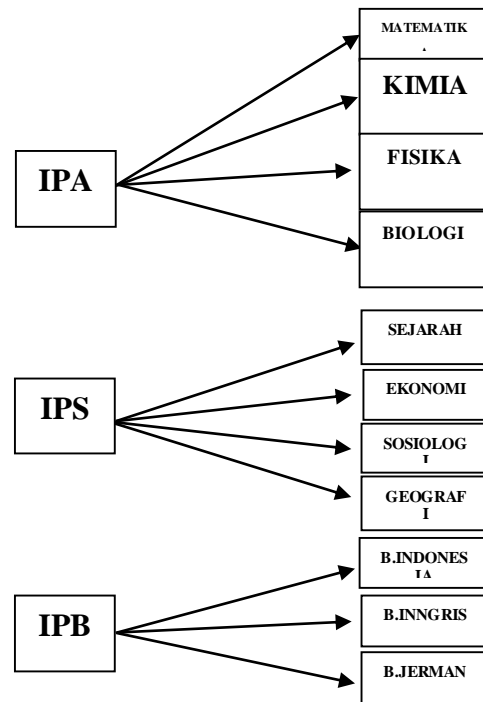
A. Kriteria Penjurusan

Pada proses pemilihan jurusan perlu diperhatikan bahwa pemilihan jurusan berdasarkan beberapa criteria sebagai berikut:

1. Pemilihan jurusan IPA berdasarkan hasil evaluasi nilai matakuliah yang menjadi ciri dari jursan IPA yaitu , MATEMATIKA, FISIKA, KIMIA, BIOLOGI.
2. Pemilihan jurusnan IPS berdasarkan hasil evaluasi nilai matapelajarn yang menjadi ciri dari jurusan IPS yaitu, GEOGRAFI, SOSIOLOGI, SEJARAH, EKONOMI.

3. Pemilihan jurusan IPB berdasarkan hasil evaluasi nilai matapelajaran yang menjadi ciri dari jurusan IPB yaitu, BAHASA INGGRIS, BAHASA INDONESIA, BAHASA JERMAN.
4. Penentuan jurusan berdasarkan nilai rata-rata dari tiap mata pelajaran yang menjadi ciri dari tiap-tiap jurusan
5. Penentuan jurusan juga berdasarkan nilai rata-rata keseluruhan mata pelajaran, apa bila nilai rata-rata keseluruhan >80, maka siswa boleh memilih jurusan yang dikehendaki, jika nilai rata-rata diantara 70-79, maka siswa boleh memilih jurusan hanya IPS dan IPB saja dan apabila nilai rata keseluruhan diantara 0 – 69 hanya dapat memilih jurusan IPB saja.

1. Korelasi jurusan dengan mata pelajaran perjurusan



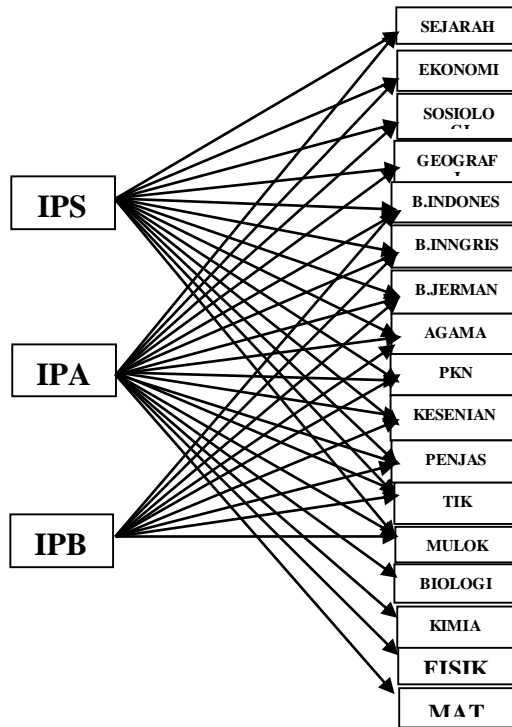
Gambar 1. Korelasi Jurusan dengan Nilai Mata Pelajaran Jurusan

B. Penerapan Fuzzy C-Means Untuk Penjurusan

Proses penentuan jurusan yang diterapkan menggunakan data hasil perhitungan rata-rata nilai mata pelajaran perjurusan dan rata-rata nilai secara keseluruhan. Nilai hasil evaluasi ini didapat setelah data dirubah menjadi data nilai dengan nilai grade atau pembobotan.

Setelah dibentuk nilai berbentuk grade/bobot maka dihitung nilai rata-rata dari tiap mata pelajaran sesuai dengan gambar korelasi mata pelajaran dibawah ini.

2. Korelasi jurusan dengan mata pelajaran keseluruhan



Gambar 2. Korelasi Jurusan dengan Seluruh Nilai Mata Pelajaran Jurusan

Dengan hasil pengkonversian nilai menjadi grade dan setelah dilakukan perhitungan rata – rata setiap mata pelajaran jurusan dan keseluruhan untuk melakukan pemilihan jurusan dengan Fuzzy C-means(FCM), penentuan penjurusan yang dilakukan menggunakan Fuzzy C-Mean, sebagai berikut:

1. Menetapkan Matriks awal dari U, berupa matrik berukuran n x m (n adalah jumlah sampel=155) dan m adalah atribut/parameter setiap data (3). X_{ij} = data sampel ke-i (i=1,2,3.....n), atribut ke-j (j=1,2,3.....,m).
2. Tentukan parameter awal sebagai berikut:
 - a. Jumlah cluster = 3

b. Pangkat = w = 2

c. Maximum Iterasi = Maxlte = 10

d. Error terkecil yang diharapkan = ϵ = 10⁻⁵

e. Fungsi objektif awal = P₀ = 0

f. Iterasi Awal = t = 1

3. Bangkitkan bilangan random μ_{ik} , i=1,2,...,n; k=1,2,...,c; sebagai elemen-elemen matriks partisi awal U.
4. Menentukan Pusat Cluster (V)

Pada bagian ini hasil sebaran dari matrik U, akan dihitung pusat cluster dari tiap-tiap cluster yang ada dengan menggunakan persamaan:

$$v_i^{(l)} = \frac{\sum_{k=1}^N \left(\mu_{ik}^{(l-1)} \right)^m z_k}{\sum_{k=1}^N \left(\mu_{ik}^{(l-1)} \right)^m}$$

5. Menghitung Fungsi Objektif Pada Iterasi Ke-1

Perhitungan Fungsi Objektif dilakukan dengan menggunakan persamaan

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left(\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right] (\mu_{ik})^w \right)$$

Periksa kondisi berhenti dengan persamaan, $|P_t - P_{t-1}| < \epsilon$, $t > \text{Maxtler}$, jika kondisi ini belum terpenuhi maka terus dilakukan iterasi sampai kondisi ini terpenuhi. Karena nilai fungsi objektif yang di dapat $P_1 = 6270,4011$ artinya kondisi $|P_t - P_{t-1}|$

$|P_t - P_{t-1}| < \epsilon$ belum terpenuhi maka dilanjutkan ke iterasi selanjutnya sampai kondisi $|P_t - P_{t-1}| < \epsilon$ terpenuhi.

Berdasarkan hasil penerapan fuzzy C-Means untuk penentuan jurusan yang ada pada SMAN 84 Jakarta, kondisi fungsi obyektif terpenuhi pada iterasi ke-4 dengan hasil sebaran derajat keanggotaan pada tiap-tiap cluster (jurusan) sebagai berikut:

1. Kelompok pertama (*cluster-1*) / Jurusan IPB akan beranggotakan siswa dengan no urut:

1, 2, 5, 11, 16, 18, 19, 21, 24, 25, 29, 30, 34, 35, 39, 43, 48, 49, 51, 52, 58, 60, 63, 64, 66, 69, 70, 71, 72, 74, 75, 81, 82, 84, 86, 88, 92, 93, 94, 95, 98, 99, 100, 103, 117, 119, 120, 122, 128, 130, 131, 132, 134, 139, 140, 146, 147, 149, 150, 155.

2. Kelompok Kedua (*cluster-2*) / Jurusan IPS akan beranggotakan siswa dengan no urut:

3, 7, 8, 9, 12, 14, 20, 22, 23, 26, 27, 36, 37, 38, 40, 41, 42, 44, 46, 47, 53, 54, 56, 59, 61, 63, 65, 67, 73, 77, 79, 80, 83, 87, 90, 96, 103, 108, 109, 111, 113, 121, 123, 124, 126, 127, 129, 133, 136, 141, 151

3. Kelompok Ketiga (*cluster-3*) / Jurusan IPA akan beranggotakan siswa dengan no urut:

4, 6, 10, 13, 15, 17, 28, 31, 32, 33, 45, 50, 55, 57, 68, 76, 78, 85, 89, 91, 97, 101, 104, 105, 106, 107, 110, 112, 114, 115, 118, 125, 135, 137, 138, 142,

143, 144, 145, 148, 152, 153, 154

KESIMPULAN

Penelitian ini dibuat untuk mengolah data nilai hasil evaluasi akademik siswa yang diperuntukan untuk menentukan pemilihan jurusan pada SMAN 84 Jakarta, dimana penentuan jurusan ini harus sesuai dengan hasil evaluasi nilai dan minat dari siswa serta pembagian kelas sesuai dengan kuota kelas jurusanyang tersedia.. Untuk menyelesaikan permasalahan ini penulis menggunakan Fuzy C-Means untuk memnentukan jurusan dan untuk mendistribusikan siswa jurusan sesuai dengan kuota kelas yang ada penulis mengguankan algoritma genetika. Kesimpulan dari penelitian ini adalah:

1. Dengan menggunakan metode FCM pihak sekolah bisa menempatkan siswa dengan jurusan yang lebih akurat.
2. Metode FCM memperoleh tingkat akurasi sebesar 72,26 % dalam penerapan penentuan jurusan pada SMA 84 Jakarta.
3. Metode Algoritma dapat mendistriusikan siswa jurusan kedalam kelas dengan perbandingan yang seimbang sesuai dengan kuota kelas yang disediakan.

UCAPAN TERIMA KASIH

Ucapan terima kasih ingin saya sampaikan kepada Allah SWT, kepada Orang Tua dan Istri Tercinta yang senantiasa mendukung baik

materil maupun moril, serta kepada Bapak Dr. Alimuddin M.Kom, selaku dosen pembimbing Thensis di Program Pasca Sarjana STMIK Nusa Mandiri

DAFTAR PUSTAKA

- Asri Yusri (2013). Aplikasi Penentuan Jurusan Mahasiswa baru Jalur PMB Dengan Algoritma Fuzzy C-Mean. *Jurnal Informatika, Sains & Teknologi*. Bekasi, 2-10.
- Luthfi Emha (2007). *Fuzzy C-Means untuk Clustering Data*. Seminar Nasional Teknologi. Jogjakarta, D-1-D7.
- Munandar Ai, Widyarto Wahyu, Harsiti (2013). Clustering Data Nilai Mahasiswa Untuk Pengelompokan kosentrasi Jurusan Menggunakan *Fuzzy Cluster Means*. Seminar Nasional Aplikasi Teknologi Informasi (SNATI). Jogjakarta, G-30 – G-33
- Novamizanti Ledyda dan Vimalakirti Retno (2010). Optimasi Logika Fuzzy Menggunakan Algoritma Genetika Pada Identifikasi Pola Tanda Tangan. Konferensi Nasional Sistem dan Informatika. Bali, 151-157.
- Sayekti Ratna, Hidayat Nurul, Soebroto Arif (2013). Implementasi Algoritma Fuzzy C-MEANS untuk Pembangunan Aturan Fuzzy Pada Pengelompokan Tingkat Risiko Penyakit Kanker Payudara. Jakarta.
- Simbolon Cary, Kusumastuti Nilamsari, Irawan Beni (2013). Clustering Lulusan Mahasiswa Matematika FMIPA UNTAN PONTIANAK Menggunakan Algoritma Fuzzy C-Means. *Buletin Ilmiah Mat.Stat dan Terapannya (Bimaster)*. Volume 2, 21-26.
- Varghese Bindhya, dkk. Clustering Student Data to Characterize Performance Patterns. *International Journal of Advance Computer Science and Application*. 138-140.
- Widyastuti Naniek, Amir Hamzah (2007). Penggunaan Algoritma Genetika dalam Peningkatan Kinerja Fuzzy Clustering Untuk Pengenalan Pola. *Berkala MIPA(2)*. Yogyakarta, 1-14.