

ANALISIS PEMBENTUKAN KELOMPOK DISKUSI PANEL SISWA MENGGUNAKAN ALGORITMA FUZZY C-MEANS DAN K-MEANS

S.S Pua Luka, I.M Candiasa, K.Y.E Aryanto

Program Studi Ilmu Komputer, Pascasarjana
Universitas Pendidikan Ganesha Singaraja Indonesia
email : czaebah@yahoo.com, candiasaimade@yahoo.co.id, yota.ernanda@undiksha.ac.id

ABSTRAK

Belajar kelompok di sekolah merupakan pembelajaran *cooperative* yang mempunyai tujuan utama agar siswa dapat bersosialisasi dan bekerjasama, terutama untuk kegiatan yang memerlukan pemecahan masalah bersama. Berdasarkan pembentukan kelompok *heterogenitas*, peneliti kemudian mengembangkan sebuah aplikasi pembentukan kelompok diskusi panel siswa SMK Negeri 1 Negara menggunakan algoritma terbaik dari hasil perbandingan algoritma *fuzzy c-means* dan *k-means*, yang diperoleh dengan membandingkan rasio simpangan baku di dalam *cluster* (s_w) terhadap rasio simpangan baku antar *cluster* (s_b). Validitas cluster optimum untuk algoritma *fuzzy c-means* diperoleh berdasarkan indeks Xie Beni dan algoritma *k-means* diperoleh berdasarkan *elbow criterion*, yaitu dengan menghitung *Root Means Square Standar Deviation* (RMSSTD) dan *R-Square* (RS). Komposisi anggota kelompok yang terbentuk dari algoritma dengan maksimal anggota sebanyak 5, antara lain: Pertama, kelas XII RPL 1 jumlah siswa 43 orang terbentuk 9 kelompok diskusi dimana 7 kelompok berjumlah 5 anggota dan 2 kelompok berjumlah 4 anggota. Kedua, kelas XII RPL 2 jumlah siswa 45 orang terbentuk 9 kelompok diskusi dengan komposisi masing-masing kelompok berjumlah 5 anggota. Ketiga, kelas XII RPL 3 jumlah siswa 46 orang terbentuk 10 kelompok diskusi dimana 6 kelompok berjumlah 5 anggota dan 4 kelompok berjumlah 4 anggota. Dan keempat, kelas XII RPL 4 jumlah siswa 43 orang terbentuk 9 kelompok dimana 7 kelompok berjumlah 5 anggota dan 2 kelompok berjumlah 4 anggota. Aplikasi pembentukan kelompok diskusi panel ini sangat efektif, dimana komposisi anggota kelompok yang langsung di ekspor ke dalam *file excel* dan ditampilkan dalam bentuk grafik akan memudahkan guru menggunakan aplikasi ini dalam melaksanakan pembelajaran *cooperative*.

Kata kunci: *Algoritma Fuzzy C-Means, Algoritma K-Means, Diskusi Panel, Pembentukan Kelompok.*

ABSTRACT

Learning groups in schools is cooperative learning that has the primary goal of enabling students to socialize and work together, especially for activities that require joint problem solving. Based on the formation of the heterogeneity group, the researcher tried to develop an application for the formation of panel discussion group of students of SMK Negeri 1 Negara using the best algorithm from the comparison of fuzzy c-means and k-means algorithms, which is obtained by comparing the standard deviation ratio in the cluster (s_w) against the standard deviation ratio between clusters (s_b). The smaller the s_w and the bigger the s_b then the method performance will get better. The optimum cluster validity for fuzzy c-means algorithm is obtained based on Xie Beni index and k-means algorithm obtained by elbow criterion, by calculating Root Means Square Standard Deviation (RMSSTD) and R-Square (RS). The composition of group members of each class is formed through application with maximum number of group members of 5, among others: First, class XII RPL 1 with the number of students 43 people formed 9 discussion groups with the composition of 7 groups of 5 members and 2 other groups amounted to 4 members. Second, class XII RPL 2 with the number of students 45 people formed 9 discussion groups with the composition of each group of 5 members. Third, class XII RPL 3 with the number of students 46 people formed 10 discussion groups with the composition of 6 groups of 5 members and 4 other groups amounted to 4 members. And fourth, class XII RPL 4 with the

number of students 43 people formed 9 groups with the composition of 7 groups amounted to 5 members and 2 other groups amounted to 4 members. The application of the formation of panel discussion groups is very effective, where the composition of group members directly exported into excel files and displayed in graphical form will make it easier for teachers to use this application in carrying out cooperative learning.

Keywords: *Fuzzy C-Means Algorithm, Group Formation, K-Means Algorithm, Panel Discussion.*

1. PENDAHULUAN

Membentuk suatu kelompok banyak mendatangkan manfaat, salah satu diantaranya adalah mendapatkan banyak informasi dari kelompok tersebut yang biasa kita kenal dengan belajar kelompok. Belajar kelompok di sekolah mempunyai tujuan utama agar siswa dapat bersosialisasi dan bekerjasama, terutama untuk kegiatan yang memerlukan pemecahan masalah bersama, seperti melakukan percobaan, berdiskusi, bermain peran, juga untuk mendorong agar siswa yang pemalu dan penakut agar mau berbicara. Siswa akan merasa aman jika berbicara dalam kelompok kecil daripada secara klasikal. Melatih siswa untuk belajar kelompok, berarti juga menyiapkan siswa tersebut untuk menjadi dewasa yang bisa bekerjasama dengan orang lain. Dalam kenyataan hidup yang membuat manusia sukses adalah kemampuannya menerapkan kecerdasan untuk bekerjasama dengan orang lain dalam mencapai tujuan bersama. Lebih-lebih dalam masyarakat modern, kemampuan bekerjasama semakin penting dan mutlak dibutuhkan [1].

Manfaat belajar dalam kelompok tertuang dalam salah satu artikel bertema *Interactive Learning Environments* dengan judul *Learning From Each Other* [2], menggambarkan betapa bagusnya suatu sesi workshop mengenai suatu desain inklusif dalam perencanaan dan penilaian pengajaran dan pembelajaran di Brighton Business School. Dalam workshop yang diikuti langsung oleh peneliti, dibentuk beberapa kelompok dengan jumlah anggota 6-7 orang membahas tentang suatu masalah tetapi tidak untuk dipresentasikan. Proses pembelajaran yang berlangsung adalah dimana masing-masing anggota dalam sesi tersebut mencatat apa yang dibahas dalam kelompoknya pada kartu yang telah tersedia, kemudian bergerak menuju kelompok lain terdekat untuk mencatat atau memperoleh minimal tiga konsep pembahasan yang berbeda dari apa yang dibahas kelompoknya. Kegiatan ini dirasakan oleh peneliti untuk benar-benar fokus terhadap apa yang disajikan dan kemudian dapat berbagi dengan yang lainnya. Manfaat dan tujuan yang diperoleh peneliti dengan diadakan sesi seperti ini adalah, mengajarkan bagaimana siswa difasilitasi untuk berpartisipasi dalam kelas dan kemudian bagaimana siswa dinilai. Pembentukan kelompok dalam workshop tersebut didasarkan pada latar belakang dan wawasan yang berbeda. Hal ini untuk menciptakan lingkungan belajar dan mengajar, pengalaman yang mendukung, memperkaya dan bisa saling menginformasikan. Lingkungan belajar yang berkontribusi mendidik dalam kegiatan itu, akan menambah pemahaman dan kualitas pengajaran dan pengalaman siswa [2].

Pembentukan kelompok yang merata dan seimbang pernah dilakukan oleh Avreghly Barrea Al-Ilman, Akbar Aditya Maulana dan Regian Siregar dari Universitas Brawijaya Malang di Program Teknologi Informasi dan Komunikasi Jurusan Informatika/Ilmu Komputer dengan judul *Sistem Pembentukan Kelompok Siswa dengan Distribusi Kemampuan Merata Menggunakan Algoritma Klusterisasi K-Means* [3], yang menjelaskan adanya suatu potensi permasalahan yang terjadi dalam proses belajar mengajar secara berkelompok. Salah satu yang disebutkan dalam penelitian tersebut adalah perbedaan kemampuan tiap kelompok yang disebabkan oleh distribusi siswa ke tiap kelompok yang random/tidak mengacu sama sekali pada kualitas siswa, sehingga berpotensi siswa yang kemampuannya tinggi berkumpul dengan yang tinggi dan yang rendah berkumpul dengan yang rendah, dan menyebabkan timpangnya hasil belajar antara kelompok tersebut. Untuk mengatasinya mereka memanfaatkan sebuah sistem pembentukan kelompok siswa dengan distribusi kemampuan merata menggunakan algoritma klusterisasi *k-means* berbasis web yang diberi nama *Grup.In*, dimana sistem ini dirancang untuk membantu pembentukan kelompok belajar siswa menggunakan algoritma klusterisasi *k-means* dengan mengacu pada nilai evaluasi pribadi tiap siswa. Pengelompokan yang dihasilkan dari aplikasi *Grup.In* ini, diperoleh beberapa kelompok belajar dengan kemampuan yang merata/seimbang [3].

Dengan kelompok belajar pola merata, tentunya guru bebas menerapkan pembelajaran *cooperative* model diskusi. Salah satu bentuk diskusi yang bisa diterapkan dengan penilaian langsung adalah diskusi panel, dimana salah satu kelompok siswa akan menjadi kelompok panelis dan yang lainnya

akan menjadi kelompok *audiens*. Selain memberi penilaian langsung, guru akan bertindak sebagai *moderator* yang mengarahkan agar diskusi panel tetap dalam masalah yang dibahas. Apakah algoritma yang digunakan dalam pembentukan kelompok diskusi tersebut adalah yang terbaik? Bagaimana analisis efektivitas kelompok diskusi panel yang terbentuk? Hal ini mendorong peneliti saat ini berusaha mengembangkan sebuah aplikasi pembentukan kelompok diskusi panel siswa berbasis Java pada SMK Negeri 1 Negara Provinsi Bali menggunakan algoritma *fuzzy c-means* dan *k-means*. Sejuahmana efektivitas komposisi anggota kelompok diskusi panel siswa yang terbentuk melalui aplikasi patut untuk dipahami. Untuk itu perlu adanya analisis terhadap aplikasi pembentukan kelompok menggunakan algoritma *fuzzy c-means* dan *k-means* tersebut.

Dalam proses analisis, persiapan awal yang dilakukan peneliti antara lain melalui tahapan; Pertama, merancang aplikasi pembentukan kelompok belajar siswa menggunakan algoritma *fuzzy c-means* dan *k-means* yang memudahkan guru dalam pelaksanaan pembelajaran *cooperative* model diskusi, salah satunya adalah diskusi panel siswa. Kedua, membangun aplikasi pembentukan kelompok belajar siswa menggunakan algoritma *fuzzy c-means* dan *k-means* dengan bahasa pemrograman GUI (*Graphical User Interface*) Java, sehingga pengimplementasiannya dalam pelaksanaan pembelajaran lebih mudah bagi guru. Ketiga, menganalisa efektivitas pembentukan kelompok belajar diskusi panel siswa menggunakan metode terbaik dari kedua algoritma, *fuzzy c-means* dan *k-means* sehingga diperoleh komposisi anggota kelompok terbaik pula dalam aplikasi tersebut.

Metode terbaik diperoleh dengan membandingkan rasio simpangan baku di dalam *cluster* (s_w) terhadap rasio simpangan baku antar *cluster* (s_b). Semakin kecil s_w dan semakin besar s_b maka kinerja metode akan semakin bagus. Jumlah *cluster* optimum dalam aplikasi yang dihasilkan dari metode terbaik juga akan dibandingkan. Validitas *cluster* optimum untuk algoritma *fuzzy c-means* diperoleh berdasarkan indeks Xie Beni dan algoritma *k-means* diperoleh berdasarkan *elbow criterion*, yaitu dengan menghitung *Root Means Square Standar Deviation* (RMSSTD) dan *R-Square* (RS). Komposisi dan jumlah *cluster* menggunakan algoritma dengan nilai s_w/s_b yang terkecil diharapkan akan digunakan sebagai pembentuk komposisi anggota kelompok belajar diskusi panel terbaik bagi siswa SMK Negeri 1 Negara Provinsi Bali.

2. KAJIAN TEORI

A. Algoritma Fuzzy C-Means

Fuzzy C-Means clustering (FCM) atau dikenal juga sebagai *fuzzy* [4], adalah salah satu metode clustering yang merupakan bagian dari metode *Hard KMeans*. FCM adalah suatu teknik pengklasteran data yang keberadaan tiap-tiap data dalam suatu *cluster* ditentukan oleh nilai/derajat keanggotaan tertentu. Teknik ini pertama kali diperkenalkan oleh Jim Bezdek pada tahun 1981 [5].

Setiap data memiliki derajat keanggotaan untuk tiap-tiap *cluster*. Dengan cara memperbaiki pusat *cluster* dan nilai keanggotaan tiap-tiap data secara berulang, maka dapat dilihat bahwa pusat *cluster* akan menuju lokasi yang tepat. Perulangan ini didasarkan pada minimasi fungsi obyektif. Fungsi obyektif yang digunakan FCM adalah sebagai berikut [5]. :

$$J(U, V; X) = \sum_{k=1}^n \sum_{i=1}^c (\mu_{ik})(d_{ik}) \quad (1)$$

$$d_{ik} = d(x_k - v_i) = \left[\sum_{j=1}^m (x_{kj} - v_{ij}) \right] \quad (2)$$

Dengan $W \in 1, \infty$, X adalah data yang dicluster :

$$X = \begin{matrix} x_{11} & \cdots & x_{1m} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nm} \end{matrix}$$

dan V adalah matrik pusat cluster:

$$V = \begin{matrix} v_{11} & \cdots & v_{1m} \\ \vdots & & \vdots \\ v_{c1} & \cdots & v_{cm} \end{matrix}$$

Nilai J_w terkecil adalah yang terbaik, sehingga:

$$J_w^*(U^*, V^*; X) = \min_{U, V} J_w(U, V, X) \quad (3)$$

Jika $d_{ik} > 0$, $\forall i, k$; $w > 1$ dan X setidaknya memiliki m elemen, maka $(U, V) \in M^m \times R^{m \times p}$ dapat meminimasi J_w hanya jika :

$$\mu_{ik} = \frac{\left[\sum_{j=1}^m (x_{ij} - v_{kj})^2 \right]^{\frac{-1}{w-1}}}{\sum_{k=1}^m \left[\sum_{j=1}^m (x_{ij} - v_{kj})^2 \right]^{\frac{-1}{w-1}}}; \quad 1 \leq i \leq m; \quad 1 \leq k \leq n \quad (4)$$

Keterangan :

1. Menentukan data yang akan di cluster X , berupa matriks berukuran $n \times m$ (n = jumlah sampel data, m = atribut setiap data). X_{ij} = data sampel ke- i ($i=1,2,\dots,n$), atribut ke- j ($j=1,2,\dots,m$).
2. Menentukan jumlah cluster (c), pangkat (w), maksimum interaksi (MaxIter), error terkecil yang diharapkan (ϵ), fungsi obyektif awal ($P_0 = 0$), iterasi awal ($t=1$).

B. Indeks Xie Beni (Index XB)

Sesuai dengan namanya indeks XB ditemukan oleh Xie dan Beni yang pertama kali dikemukakan pada tahun 1991 [6]. Validitas dalam *fuzzy c-means* ditentukan oleh banyak kelompok optimum melalui perhitungan indeks validitas. Indeks ini bertujuan untuk mengukur rasio dari total variasi dalam *cluster* (S) dan pemisah antar *cluster* (d_{min}) [6]. Banyaknya *cluster* optimum dengan meminimalkan nilai indeks [7]. Kriteria banyak kelompok optimum diberikan oleh nilai XB yang minimum pada lembah pertama [8]. Rumus *index XB* :

$$XB = \frac{\sum_{i=1}^c \sum_{j=1}^n \mu_{ik}^w * \|V_i - X_j\|^2}{n * \min_{i,j} \|V_i - V_j\|^2} \quad (5)$$

Dengan c menyatakan banyak *cluster*, μ_{ik} adalah tingkat keanggotaan, $\|V_i - X_j\|^2$ adalah jarak observasi (X_j) dengan pusat *cluster*, V_i adalah pusat *cluster*, n merupakan banyak objek yang akan dikelompokkan, $\min_{i,j} \|V_i - X_j\|^2$ menyatakan jarak minimum antara pusat *cluster* V_i dan V_j . Rekomendasi untuk menggunakan indeks XB tertuang dalam penelitian Duo tahun 2007, yang menyatakan bahwa indeks XB memiliki ketepatan dan keandalan yang tinggi baik untuk memberikan banyak kelompok optimum pada metode hard partition seperti *k-means cluster* maupun pada *fuzzy c-means cluster* [9].

C. Algoritma K-Means

Algoritma *K-Means* merupakan algoritma yang membutuhkan parameter input sebanyak k dan membagi sekumpulan n objek kedalam k *cluster* sehingga tingkat kemiripan antar anggota dalam satu *cluster* tinggi sedangkan tingkat kemiripan dengan anggota pada *cluster* lain sangat rendah. Kemiripan anggota terhadap *cluster* diukur dengan kedekatan objek terhadap nilai *mean* pada *cluster* atau dapat disebut sebagai *centroid cluster* atau pusat massa [10].

Langkah-langkah untuk *k-means* berarti algoritma yang diberikan di bawah [11]:

1. Memilih secara acak vector masukan K (titik data) untuk menginisialisasi *cluster*.
2. Pencarian terdekat : untuk setiap *vector input*, menemukan pusat *cluster* yang paling dekat, dan menetapkan bahwa *vector* memasukan *cluster* yang sesuai.
3. *Update* pusat *cluster* disetiap *cluster* dengan menggunakan rata-rata (*centroid*) dari *vector* untuk *cluster* itu.

$$c_k = \left(\frac{1}{n_k} \right) \sum d_i \quad (6)$$

4. Ulangi langkah 2 dan 3 sampai tidak ada lagi perubahan nilai yang lebih dalam.

D. Elbow Criterion (RMSSTD dan RS)

Dalam melakukan proses analisis *cluster*, pengujian atas kevalidan atau kesahihan suatu hasil analisis *cluster* terdapat dua cara [12], yaitu pertama *internal test*, suatu cara pengujian dengan membandingkan hasil *cluster* yang terbentuk dari beberapa metode berbeda yang digunakan; Kedua solusi *cluster* yang diajukan oleh Sharma. Tahapan validasi dalam analisis *cluster* yang dilakukan oleh Sharma untuk menguji apakah *cluster* yang terbentuk dari hasil subjektifitas peneliti telah valid atau tidak. Uji validasi terhadap *cluster* yang terbentuk dilihat dari plot nilai RMSSTD (*Root Mean Square Standard Deviation*) dan nilai RS (*R-Square*) terhadap jumlah *cluster*. RMSSTD merupakan ukuran homogenitas *cluster* baru, dengan syarat nilai harus kecil, sedangkan RS merupakan ukuran heterogenitas *cluster*, dengan syarat nilai harus besar [12]. RMSSTD dihitung menggunakan rumus sebagai berikut :

$$RMSSTD = \sqrt{\frac{\sum_{i=1}^k \sum_{j=1}^d (\sum_{k=1}^{N_i} (x_{kj} - \mu_j)^2)}{\sum_{i=1}^k \sum_{j=1}^d (N_i - 1)}} \quad (7)$$

RS dihitung dengan rumus sebagai berikut :

$$RS = \frac{SS_t - SS_w}{SS_t} \quad (8)$$

dimana :

$$SS_t = \sum_{j=1}^d \left(\sum_{k=1}^N (x_{kj} - \mu_j)^2 \right) \quad (9)$$

dan

$$SS_w = \sum_{i=1}^k \sum_{j=1}^d \left(\sum_{k=1}^{N_i} (x_{kj} - \mu_j)^2 \right) \quad (10)$$

Notasi :

- x_{kj} : data ke-k yang ada di dalam *cluster* untuk dimensi ke-j
- μ_j : means / rata-rata nilai dari variabel dimensi ke-j
- N_i : jumlah data di dalam *cluster* ke-i
- N : jumlah data keseluruhan
- d : jumlah dimensi dari data
- k : jumlah cluster

Presentasi hasil perbandingan plot nilai RMSSTD dan RS yang akan membentuk siku (*elbow*) pada suatu titik menjadi salah satu kriteria dalam menentukan jumlah *cluster* terbaik (*optimum*) untuk algoritma *k-means*.

3. METODE

A. Data yang Digunakan

Data yang digunakan dalam penelitian ini adalah data yang diperoleh langsung dari hasil penilaian belajar siswa sebelumnya berupa Penilaian Akhir Semester (PAS) ganjil, Penilaian Akhir Semester (PAS) genap, Penilaian Keterampilan (PK) dan Penilaian Sikap (PS) serta hasil *pretest* siswa bersangkutan yang diperoleh pada awal dilakukan pembelajaran tahun pelajaran. Data nilai tersebut diperoleh dari empat (4) kelas XII Rekeyasa Perangkat Lunak. Data penilaian siswa salah satu kelas, antara lain :

Tabel 1. Data dan Nilai Siswa

NO	NIS	NAMA SISWA	JENIS KELAMIN	PAS GANJIL	PAS GENAP	PRETEST	PK SISWA	PS SISWA
1	AA01	AAAAA	Laki-laki	85	85.75	85	86	87
2	AA02	BBBBB	Laki-laki	81	82.75	82	83	83

NO	NIS	NAMA SISWA	JENIS KELAMIN	PAS GANJIL	PAS GENAP	PRETEST	PK SISWA	PS SISWA
3	AA03	CCCCC	Laki-laki	85	85.25	84	86	86
4	AA04	DDDDD	Laki-laki	81	83.5	80	84	85
5	AA05	EEEEEE	Laki-laki	85	83.25	82	84	83
6	AA06	FFFFFF	Perempuan	85	83.25	82	84	83
7	AA07	GGGGG	Laki-laki	85	84.5	83	85	85
8	AA08	HHHHH	Perempuan	85	85.5	84	86	86
9	AA09	IIIII	Perempuan	81	83.5	82	84	83
10	AA10	JJJJJ	Perempuan	82	84.25	82	85	85
11	AA11	KKKKK	Perempuan	85	85.5	83	86	86
12	AA12	LLLLL	Perempuan	83	85.25	83	86	86
13	AA13	MMMMM	Perempuan	83	84.5	83	85	85
14	AA14	NNNNN	Laki-laki	83	84.5	82	85	85
15	AA15	OOOOO	Laki-laki	81	83.5	82	84	83
16	AA16	PPPPP	Perempuan	81	84.25	83	85	84
17	AA17	QQQQQ	Perempuan	81	83.5	83	84	83
18	AA18	RRRRR	Perempuan	82	84.25	83	85	84
19	AA19	SSSSS	Perempuan	82	84.25	83	85	84
20	AA20	TTTTT	Perempuan	83	85.25	85	86	85
21	AA21	UUUUU	Perempuan	84	83.25	82	84	83
22	AA22	VVVVV	Laki-laki	85	84.25	83	85	85
23	AA23	WWWWW	Laki-laki	85	83.25	82	84	83
24	AA24	XXXXX	Perempuan	82	84.25	83	85	84
25	AA25	YYYYY	Laki-laki	82	84.25	83	85	84
26	AA26	ZZZZZ	Laki-laki	83	83.25	82	84	83
27	AA27	AAABB	Laki-laki	83	84.5	83	85	85
28	AA28	AAACC	Laki-laki	82	84.25	83	85	84
29	AA29	AAADD	Perempuan	83	84.5	83	85	84
30	AA30	AAEEE	Perempuan	83	84.5	83	85	84
31	AA31	AAAFF	Laki-laki	85	85.5	84	86	85
32	AA32	AAAGG	Perempuan	83	84.5	83	85	84
33	AA33	AAHHH	Laki-laki	81	83.5	83	84	83
34	AA34	AAIIL	Perempuan	83	84.25	83	85	84
35	AA35	AAAJJ	Laki-laki	83	84	83	84	84
36	AA36	AAAKK	Perempuan	85	85.5	84	86	85
37	AA37	AAALL	Laki-laki	83	84.5	83	85	84
38	AA38	AAAMM	Laki-laki	83	84.5	83	85	84
39	AA39	AAANN	Laki-laki	84	84.5	83	85	85
40	AA40	AAAOO	Laki-laki	83	84.5	83	85	84
41	AA41	AAAPP	Laki-laki	83	84.25	83	85	84
42	AA42	AAAQQ	Perempuan	83	84.25	83	85	84
43	AA43	AAARR	Laki-laki	82	84.25	83	85	84

B. Proses dalam Aplikasi Pembentukan Kelompok Diskusi

Tahapan proses pembentukan kelompok diskusi dalam aplikasi untuk masing-masing kelas siswa, antara lain:

1. Proses Data Master

Dalam data master ini dilakukan proses penginputan data jurusan, kelas, guru dan mata pelajaran. Data-data tersebut akan ditampilkan kembali ke tampilan awal aplikasi (*Home*) bersama dengan identitas sekolah.

2. Proses Pengclusteran

Setelah proses data master selesai, proses pengclusteran dalam aplikasi dapat dilakukan sesuai kelas siswa yang dipilih. Ini berarti setiap terjadi pergantian kelas, perlu dilakukan proses identitas data personal kembali. Tahapan dalam proses pengclusteran, antara lain:

a. Input data siswa dan penilaian siswa

Input data siswa dan penilaian siswa yang dapat dilakukan secara manual atau dengan *import* data dan penilaian siswa tersebut sesuai masing-masing kelas dengan *format file excel*. Setelah data siswa dan penilaian siswa tersimpan ke *database*, langkah selanjutnya adalah proses pengclusteran menggunakan algoritma *fuzzy c-means* dan *k-means*.

b. Proses Pengclusteran menggunakan *Fuzzy C-Means* (FCM) dan *K-Means* (KM) :

(a) Pengclusteran dengan *Fuzzy C-Means* (FCM)

Clusterisasi siswa masing-masing kelas dengan *fuzzy c-means*, diawali dengan menginputkan maksimal proses iterasi dan jumlah *cluster* yang diinginkan. Jumlah *cluster* dalam aplikasi ini ditetapkan sebanyak 4 *cluster*, 3 *cluster* dan 2 *cluster*. Masing-masing jumlah *cluster* tersebut akan diterapkan ke masing-masing kelas siswa agar memperoleh jumlah *cluster* terbaik untuk pembentukan kelompok diskusi.

Iterasi masing-masing jumlah *cluster* dengan *fuzzy c-means* ini akan berhenti ketika mencapai maksimal iterasi yang diinput sebelumnya atau $p < \text{Error Terkecil} (\xi)$, dimana ξ yang digunakan dalam aplikasi sebesar 0,01 ($1/10^2$), $p = (p_i - p_{(i-1)})$ dan p_i adalah total fungsi objektif masing-masing iterasi.

(b) Pengclusteran dengan *K-Means* (KM)

Clusterisasi dengan *k-means* masing-masing kelas siswa, diawali dengan menginputkan jumlah *centroid* yang diinginkan. Jumlah *centroid* dalam aplikasi ini juga ditetapkan sebanyak 4 *centroid*, 3 *centroid* dan 2 *centroid*, yang masing-masing akan diterapkan ke masing-masing kelas siswa agar memperoleh jumlah *cluster* terbaik untuk pembentukan kelompok diskusi.

Iterasi *cluster* dengan *k-means* akan berhenti ketika hasil cluster iterasi terakhir dengan iterasi sebelumnya sudah tidak ada perbedaan.

3. Proses Penentuan Metode dan Jumlah *Cluster* Terbaik

Penentuan metode terbaik dari algoritma *fuzzy c-means* dan *k-means* adalah dengan membandingkan rasio simpangan baku di dalam *cluster* (s_w) terhadap rasio simpangan baku antar *cluster* (s_b) iterasi terakhir masing-masing kedua algoritma tersebut. Algoritma dengan (s_w/s_b) terkecil akan ditentukan sebagai metode terbaik dalam aplikasi.

Jumlah *cluster* terbaik dalam aplikasi ini diperoleh dengan membandingkan jumlah *cluster* optimum iterasi terakhir dari metode terbaik. Validitas *cluster* optimum dihitung berdasarkan indeks Xie Beni untuk algoritma *fuzzy c-means* dan *elbow criterion* (RMSSTD dan RS) untuk algoritma *k-means*.

4. Proses Pembentukan Kelompok Diskusi

Dalam aplikasi, proses pembentukan kelompok diskusi dilakukan berdasarkan jumlah dan komposisi *cluster* dari algoritma terbaik. Tahapan pembentukan kelompok, antara lain :

a. Menentukan urutan *cluster* terbaik dari algoritma terbaik dengan cara menghitung rata-rata masing-masing pusat *cluster*. *Cluster* dengan rata-rata terbesar merupakan *cluster* terbaik.

b. Menentukan maksimal anggota kelompok. Dalam aplikasi ini maksimal anggota dalam 1 kelompok diskusi dapat dipilih sebanyak 6 anggota, 5 anggota dan 4 anggota. Jumlah kelompok dan komposisi anggota kelompok akan sangat tergantung dengan penentuan maksimal anggota kelompok ini.

c. Melakukan pemetaan (*mapping*) kelompok diskusi.

Pemetaan kelompok diskusi akan disesuaikan dengan jumlah maksimal anggota yang dipilih sebelumnya.

d. Menampilkan grafik hasil pembentukan kelompok dan ekspor *file* hasil dalam format *excel*

4. HASIL DAN PEMBAHASAN

Pada penelitian ini membahas bagaimana proses dalam aplikasi pembentukan kelompok diskusi menggunakan algoritma *fuzzy c-means* dan *k-means* serta proses validasi penentuan algoritma dan jumlah cluster terbaik untuk pembentukan kelompok diskusi dalam aplikasi tersebut.

A. Pengclustering Algoritma *Fuzzy C-Means* dan *K-Means* dalam Aplikasi Pembentukan Kelompok

Hasil pengclustering yang diperoleh peneliti dari aplikasi pembentukan kelompok menggunakan algoritma *fuzzy c-means* dan *k-means* adalah sebagai berikut :

Tabel 2. Rekapitulasi Hasil Pengclustering Aplikasi

Kelas : XII RPL 1
 Jumlah Siswa : 43 Orang
 Metode : *Algoritma Fuzzy C-Means (FCM)*

No	Jumlah Cluster	Cluster				Iterasi ke-	S _w	S _b	S _w /S _b	Xie Beni
		I	II	III	IV					
1	4	19	7	8	9	14	0.727957	0.193838	3.755492	1.747478
2	3	19	6	18	-	9	0.774689	0.05742	13.491623	1.839857
3	2	21	22	-	-	9	0.77431	0.042166	18.363373	1.799797

Metode : *Algoritma K-Means (KM)*

No	Jumlah Centroid	Cluster				Iterasi ke-	S _w	S _b	S _w /S _b	RMSSTD	RS
		I	II	III	IV						
1	4	8	14	5	16	4	0.353310	0.847045	0.417109	1.016278	0
2	3	12	16	15	-	4	0.527759	0.691433	0.763283	1.003494	0
3	2	11	32	-	-	5	0.448367	1.006723	0.445373	0.991181	0

Kelas : XII RPL 2
 Jumlah Siswa : 45 Orang
 Metode : *Algoritma Fuzzy C-Means (FCM)*

No	Jumlah Cluster	Cluster				Iterasi ke-	S _w	S _b	S _w /S _b	Xie Beni
		I	II	III	IV					
1	4	24	2	10	9	9	0.737881	0.181659	4.061902	7.507180
2	3	18	4	23	-	10	0.992963	0.248336	3.998466	11.612735
3	2	24	21	-	-	9	0.894315	0.044621	20.042469	6.634845

Metode : *Algoritma K-Means (KM)*

No	Jumlah Centroid	Cluster				Iterasi ke-	S _w	S _b	S _w /S _b	RMSSTD	RS
		I	II	III	IV						
1	4	19	16	2	8	5	0.261171	0.982686	0.265773	0.676252	0
2	3	22	15	8	-	7	0.356748	1.106625	0.322375	0.668153	0
3	2	28	17	-	-	3	0.525897	1.053381	0.499247	0.660338	0

Kelas : XII RPL 3
 Jumlah Siswa : 46 Orang
 Metode : *Algoritma Fuzzy C-Means (FCM)*

No	Jumlah Cluster	Cluster				Iterasi ke-	S _w	S _b	S _w /S _b	Xie Beni
		I	II	III	IV					
1	4	22	0	14	10	12	1.206795	0.165493	7.292121	9.223E ¹²
2	3	16	7	23	-	10	1.170391	0.237691	4.924002	4.107470
3	2	21	25	-	-	12	1.152303	0.114280	10.083155	4.348895

Metode : *Algoritma K-Means (KM)*

No	Jumlah Centroid	Cluster				Iterasi ke-	Sw	Sb	Sw/Sb	RMSSTD	RS
		I	II	III	IV						
1	4	13	13	20	0	3	0.529643	1.166331	0.45411	1.190173	0
2	3	13	13	20	-	3	0.529643	1.166331	0.45411	1.190173	0
3	2	13	33	-	-	2	0.810495	1.222438	0.663015	1.176571	0

Kelas : XII RPL 4
 Jumlah Siswa : 43 Orang
 Metode : *Algoritma Fuzzy C-Means (FCM)*

No	Jumlah Cluster	Cluster				Iterasi ke-	Sw	Sb	Sw/Sb	Xie Beni
		I	II	III	IV					
1	4	20	2	14	7	14	0.632304	0.170056	3.718210	0.571867
2	3	8	9	26	-	9	0.607094	0.144115	4.212566	1.378251
3	2	19	24	-	-	11	0.761854	0.067823	11.232974	1.340109

Metode : *Algoritma K-Means (KM)*

No	Jumlah Centroid	Cluster				Iterasi ke-	Sw	Sb	Sw/Sb	RMSSTD	RS
		I	II	III	IV						
1	4	11	4	24	4	6	0.640789	0.946835	0.676769	1.670062	0
2	3	23	16	4	-	4	0.756920	0.970192	0.780175	1.649054	0
3	2	39	4	-	-	3	1.003424	1.157461	0.866918	1.628819	0

B. Validitas

1) Hasil Perbandingan Algoritma dengan Sw/Sb Terkecil dalam Aplikasi Pembentukan Kelompok

Hasil perbandingan algoritma *fuzzy c-means* dan *k-means* dalam aplikasi pembentukan kelompok diskusi yang divalidasi berdasarkan nilai Sw/Sb terkecil adalah sebagai berikut :

Tabel 3. Rekapitulasi Hasil Perbandingan Metode

No	Kelas	Metode sw/sb Terkecil	Jumlah Cluster	XB/RMSSTD	RS	Urutan Cluster
1	XII RPL 1	Algoritma <i>K-Means (KM)</i>	4	1.016278	0	
			3	1.003494	0	
			2	0.991181	0	1, 2
2	XII RPL 2	Algoritma <i>K-Means (KM)</i>	4	0.676252	0	
			3	0.668153	0	
			2	0.660338	0	1, 2
3	XII RPL 3	Algoritma <i>K-Means (KM)</i>	4	1.190173	0	
			3	1.190173	0	
			2	1.176571	0	2, 1
4	XII RPL 4	Algoritma <i>K-Means (KM)</i>	4	1.670062	0	
			3	1.649054	0	
			2	1.628819	0	1, 2

2) Hasil Pemetaan Pembentukan Kelompok dalam Aplikasi dengan Metode Terbaik

Hasil pemetaan pembentukan kelompok dari algoritma dan jumlah *cluster* terbaik adalah sebagai berikut :

Tabel 4. Rekapitulasi Hasil Pemetaan Pembentukan Kelompok dengan Metode Terbaik

No	Kelas Siswa	Jumlah Siswa	Jumlah Kelompok	Komposisi Anggota
1	XII RPL 1	43	9	7 kelompok = 5 anggota 2 kelompok = 4 anggota
2	XII RPL 2	45	9	9 kelompok = 5 anggota
3	XII RPL 3	46	10	6 kelompok = 5 anggota 4 kelompok = 4 anggota
4	XII RPL 4	43	9	7 kelompok = 5 anggota 2 kelompok = 4 anggota

Implementasi aplikasi pembentukan kelompok diskusi panel menggunakan algoritma *fuzzy c-means* dan *k-means* berbasis Java, pada masing-masing kelas XII Rekayasa Perangkat Lunak (RPL) SMK Negeri 1 Negara sebanyak 4 kelas diperoleh algoritma terbaik dan jumlah *cluster* optimum. Algoritma terbaik ditentukan berdasarkan perbandingan rasio simpangan baku di dalam *cluster* (s_w) terhadap rasio simpangan baku antar *cluster* (s_b) terkecil, sedangkan jumlah cluster diperoleh dari nilai minimum indeks Xie Beni (XB) untuk algoritma *fuzzy c-means* dan perbandingan prosentase perubahan kedua nilai (RMSSD dan RS) untuk algoritma *k-means*. Hasil perbandingan metode dalam aplikasi diperoleh metode terbaik adalah algoritma *k-means* untuk masing-masing kelas dengan jumlah cluster sebanyak 2 *cluster*. Urutan *cluster* optimum mulai dari yang terbaik dapat dilihat pada table 3.

Selain jumlah dan urutan *cluster* optimum dari metode terbaik digunakan sebagai penentuan komposisi anggota kelompok diskusi panel pada masing-masing kelas, jumlah maksimal anggota dalam kelompok juga sangat berperan penting dalam penentuan komposisi anggota kelompok (table 4). Dari komposisi anggota yang terbentuk, pemetaan kelompok sebagai hasil akhir pembentukan kelompok diskusi kemudian diproses dengan memberikan penomoran berdasarkan urutan susunan anggota *cluster* dari yang terbaik sesuai jumlah kelompok dan jumlah siswa masing-masing kelas. Hal ini dilakukan agar komposisi anggota dalam kelompok siswa menjadi bervariasi, terdiri dari siswa pintar, biasa dan kurang. Kelompok siswa yang terbentuk melalui aplikasi akan langsung di *ekspor* ke dalam *file excel* dan ditampilkan dalam bentuk grafik.

5. SIMPULAN

Metode terbaik yang diperoleh berdasarkan perbandingan rasio simpangan baku di dalam *cluster* (s_w) terhadap rasio simpangan baku antar *cluster* (s_b) terkecil yaitu algoritma *k-means*, diperoleh melalui aplikasi pembentukan kelompok diskusi panel siswa SMK Negeri 1 Negara menggunakan algoritma *fuzzy c-means* dan *k-means* berbasis Java, dengan jumlah *cluster* optimum masing-masing kelas sebanyak 2 *cluster* yang diperoleh berdasarkan *elbow criterion*, yaitu dengan menghitung *Root Means Square Standar Deviation* (RMSSTD) dan *R-Square* (RS). Urutan *cluster* optimum terbaik yang diperoleh berdasarkan maksimum rata-rata pusat clusternya antara lain : pertama, untuk kelas XII RPL 1 adalah *cluster* 1 kemudian *cluster* 2. Kedua, untuk kelas XII RPL 2 adalah *cluster* 1 kemudian *cluster* 2. Ketiga, untuk kelas XII RPL 3 adalah *cluster* 2 kemudian *cluster* 1. Dan keempat, untuk kelas XII RPL 4 adalah *cluster* 1 kemudian *cluster* 1. Komposisi dan jumlah *cluster* untuk jumlah maksimal anggota kelompok sebanyak 5, adalah : pertama, kelas XII RPL 1, jumlah siswa 43 orang terbentuk 9 kelompok diskusi dengan komposisi 7 kelompok berjumlah 5 anggota dan 2 kelompok lainnya berjumlah 4 anggota. Kedua, kelas XII RPL 2 dengan jumlah siswa 45 orang terbentuk 9 kelompok diskusi dengan komposisi masing-masing kelompok berjumlah 5 anggota. Ketiga, kelas XII RPL 3 dengan jumlah siswa 46 orang terbentuk 10 kelompok diskusi dengan komposisi 6 kelompok berjumlah 5 anggota dan 4 kelompok lainnya berjumlah 4 anggota. Dan keempat, kelas XII RPL 4 dengan jumlah siswa 43 orang terbentuk 9 kelompok dengan komposisi 7 kelompok berjumlah 5 anggota dan 2 kelompok lainnya berjumlah 4 anggota. Aplikasi pembentukan kelompok diskusi panel menggunakan algoritma *fuzzy c-means* dan *k-means* berbasis Java ini sangat efektif untuk diterapkan dimana komposisi anggota kelompok yang

langsung di eksport ke dalam *file* excel dan ditampilkan dalam bentuk grafik, memudahkan bagi guru yang menggunakan aplikasi ini dalam melaksanakan pembelajaran *cooperative*.

UCAPAN TERIMAKASIH

Terimakasih yang sebesar-besarnya disampaikan kepada bapak Kepala Sekolah dan seluruh staf manajemen SMK Negeri 1 Negara Provinsi Bali yang telah memberikan ijin sebagai tempat pelaksanaan penelitian dan juga data-data yang diperlukan sehingga penelitian ini dapat terselesaikan dengan baik.

DAFTAR PUSTAKA

- [1] Schmuck (Eds.), *Learning to Cooperate, Cooperating to Learn*, NewYork: Plenum, 1985.
- [2] Pericles Asher Rospigliosi, "Learning from Each Other," *J. Internasional., Interactive Learning Environments, Routledge Taylor and Francis Group*.
- [3] Avregly Barra Al-Ilman, Akbar Aditya Maulana dan Regian Siregar, "Sistem Pembentukan Kelompok Siswa dengan Distribusi Kemampuan Merata Menggunakan Algoritma Klusterisasi K-Means," *J.Fibonacci.*, vol.1, no.2..
- [4] Z. Hilal Inana and Mehmet Kuntalp, "A Study on Fuzzy C-Means Clustering-Based Systems in Automatic Spike Detection," pp. 1160-1166.
- [5] Kusumadewi, S dan Purnomo, *Aplikasi Logika Fuzzy untuk Pendukung Keputusan*, Yogyakarta: Graha Ilmu, 2004.
- [6] Xie, X.I dan Beni, G, "A Validity Measure for Fuzzy Clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence.*, vol. 13, pp. 841-847.
- [7] Jansen, S.M.H, *Customer Segmentation and Customer Profilling for a mobile Telecommunications Company Based on Usage Behavior : A vodafone Case Study*, Maastricht: University of Maastricht, 2007.
- [8] Kuo, L.W. dan Miin, S.Y., "A Cluster Validity Index for Fuzzy Clustering," dalam *Pattern Recognition Letter.* 26 , pp. 1275-1291.
- [9] Duo, C., Xue, L. dan Du-Wu, C., ".An Adaptive Cluster Validity Index for the Fuzzy C-Means," *International Journal of Computer Science and Network Security*, vol. 7, no. 2, pp. 146-156.
- [10] Widyawati, N., *Perbandingan Clustering Based on Frequent Word Sequence (CFWS) dan K-Means untuk Pengelompokan Dokumen Berbahasa Indonesia.*, Bandung: Fakultas Pendidikan Matematika dan Ilmu Pengetahuan Alam Universitas Pendidikan Indonesia, 2010.
- [11] Prasetyo, E, *Data Mining: Konsep dan Aplikasi Menggunakan Matlab*, Yogyakarta: Andi Offset, 2012.
- [12] Sharma, S, *Applied Multivariate Technique*, Canada: John Wiley&Sons, 1996.
- [13] Aditya Ramadhan, Zuliar Efendi dan Mustakim, "Perbandingan K-Means dan Fuzzy C-Means untuk Pengelompokan Data User Knowledge Modeling", *Seminar Nasional Teknologi Informasi, Komunikasi dan Industri (SNTIKI)*.
- [14] Ni Putu Eka Merliana, Ernawati dan Alb. Joko Santoso, "Analisa Penentuan Jumlah Cluster Terbaik pada Metode K-Means Clustering", *J. Prosiding.*, *Seminar Nasional Multi Disiplin Ilmu & Call For Papers Unisbank (SENDI_U)*.
- [15] Sarita Budiyani Purnamasari, Hasbi Yasin dan Triastuti Wuryandari, "Pemilihan Cluster Optimum pada Fuzzy C-Means (Studi Kasus : Pengelompokan Kabupaten/Kota Provinsi Jawa Tengah Berdasarkan Indikator Indeks Pembangunan Manusia)", *J. Gaussian.*, vol. 3, no. 3.