

IMPLEMENTASI ALGORITMA K-MEANS UNTUK MENENTUKAN KELOMPOK PENGAYAAN MATERI MATA PELAJARAN UJIAN NASIONAL (STUDI KASUS: SMP NEGERI 101 JAKARTA)

Fenty Eka M. Agustin¹, Ardini Fitria², Anif Hanifah S³

^{1,2,3}Program Studi Teknik Informatika, Fakultas Sains dan Teknologi
Universitas Islam Negeri Syarif Hidayatullah

¹fenty.eka@uinjkt.ac.id, ²ardiniifitria@gmail.com, ³anifhanifa@yahoo.com

ABSTRAK

Pengayaan materi merupakan salah satu persiapan peserta didik untuk menghadapi Ujian Nasional. Di SMP Negeri 101 Jakarta terdapat dua pengayaan materi, yaitu pengayaan materi wajib dan pengayaan materi khusus. Pengayaan materi khusus dilaksanakan dengan melihat hasil akhir rapor semester 5. Proses pengelompokkan kemampuan siswa untuk melaksanakan pengayaan materi khusus masih belum maksimal karena kemampuan siswa tersebut tidak hanya diukur dari rapor terakhir saja, melainkan nilai rapor semester 1 hingga 5 berikut nilai tes terakhir untuk menambah keakuratan data. Untuk itu diperlukan solusi yang dapat mengatasi kesulitan tersebut. Metode *clustering* dengan menggunakan algoritma K-Means diimplementasikan dalam aplikasi ini. Jumlah *cluster* ada empat sesuai jumlah mata pelajaran UN, sedangkan jumlah sampel data adalah 12 siswa yang memiliki nilai terendah. Aplikasi ini selain menampilkan pengelompokkan kemampuan siswa pada mata pelajaran Ujian Nasional, juga dapat digunakan untuk memantau perkembangan kemampuan setelah mengikuti pengayaan materi.

Kata kunci : *Ujian Nasional, Clustering, Algoritma K-Means*

I. PENDAHULUAN

Menurut UU Nomor 20 tahun 2003 pasal 58 tentang Sistem Pendidikan Nasional menyebutkan bahwa terdapat dua jenis evaluasi, yaitu: evaluasi internal yang dilakukan oleh pendidik terhadap dua jenis evaluasi, yaitu evaluasi internal yang dilakukan pendidik terhadap hasil belajar peserta didik, dan evaluasi eksternal yang dilakukan oleh lembaga negara terhadap peserta didik, satuan pendidikan dan program pendidikan yang lebih dikenal dengan Ujian Nasional yang menjadi standar dalam penentuan kelulusan siswa dasar, menengah pertama, dan juga menengah umum.

Pentingnya pengayaan materi ditegaskan oleh Kepala Dinas Pendidikan DKI Jakarta Lasro Marbun bahwa hal tersebut dilaksanakan sebagai persiapan peserta didik untuk menghadapi Ujian Nasional [9]. Sementara itu, dalam jurnal Efektifitas Bimbingan Belajar Menghadapi Ujian Nasional Pelajaran Sosiologi SMA Muhammadiyah 2 Pontianak, terdapat pengaruh signifikan pengayaan materi terhadap keberhasilan siswa mencapai Ujian Nasional, bahwa makin efektif bimbingan belajar, maka makin tinggi pencapaian Ujian Nasional siswa pada mata pelajaran tersebut [5].

Salah satu sekolah yang melakukan pengayaan materi adalah SMP Negeri 101 Jakarta. Hasil wawancara dengan Nurlailah selaku Wakil Kepala Sekolah Bidang Kurikulum, bahwa di SMP Negeri

101 Jakarta terdapat dua pengayaan materi, yaitu pengayaan materi wajib dan pengayaan materi khusus. Pengayaan materi wajib dilakukan terhadap seluruh siswa berdasarkan kelas. Sementara pengayaan materi khusus dilaksanakan dengan merujuk hasil akhir rapor siswa semester 5.

Adapun pengelompokkan kelas pengayaan materi khusus berdasarkan nilai terendah dari mata pelajaran Ujian Nasional (IPA, Matematika, Bahasa Indonesia, dan Bahasa Inggris) dalam rapor semester 5. Jika siswa mendapatkan nilai terendah pada mata pelajaran tertentu, maka siswa mengikuti pengayaan materi khusus pada mata pelajaran tersebut.

Merujuk penjelasan Nurlailah, sejauh ini terdapat kelemahan pada pengelompokkan siswa dengan menggunakan nilai rapor semester 5 saja. Dampaknya, hasil pengelompokkan belum akurat berdasarkan kemampuan siswa. Hal ini karena materi UN mencakup materi dari materi kelas VII atau mulai semester 1. Oleh karena itu, guru juga harus mempertimbangkan nilai rapor semester 1 sampai dengan 5 dan nilai tes terakhir sebagai acuan untuk proses pengelompokkan siswa. Proses ini tidaklah mudah mengingat sekolah belum memiliki database yang dapat diolah menjadi informasi kemampuan akademik siswa. Nilai rapor hanya menginformasikan capaian akademik tiap semester saja.

Kebutuhan untuk memberikan tambahan materi belajar atau yang biasa disebut pengayaan materi bagi siswa kelas IX, membuat pihak sekolah harus

melakukan penggalian data (*datamining*). Salah satu informasi yang harus digali berdasarkan database nilai adalah penguasaan siswa terhadap materi yang akan diujikan dalam ujian nasional (UN). UN akan mengujikan materi Matematika, IPA, Bahasa Indonesia, dan Bahasa Inggris. Sekolah memberi kebijakan untuk mengadakan kelas pengayaan materi khusus untuk memperdalam penguasaan materi UN. Saat ini pengayaan materi hanya siswa dibagi dalam kelompok berdasarkan nilai yang diperoleh pada semester 5 saja secara pareto. Supaya tujuan kelas pengayaan materi ini tercapai, yaitu untuk memperkuat penguasaan materi UN. Maka dibuatlah sistem pendukung pengambilan keputusan, untuk membantu pengelompokan siswa yang cepat dan akurat. Proses pembuatan sistem didukung oleh penggunaan algoritma K-Means.

Algoritma K-Means merupakan salah satu algoritma populer yang digunakan dalam proses *clustering* dataset karena kesederhanaan algoritmanya [1]. Selain mampu menangani dataset dengan ukuran yang besar, algoritma ini juga memiliki tiga keunggulan yaitu kompleksitas waktu, kompleksitas ruang penyimpanan, dan pemrosesan tidak bergantung pada urutan centroid yang digunakan [7]. Disamping itu, algoritma K-Means juga bersifat *versatile*, yaitu mudah melakukan modifikasi di setiap tahapan dalam algoritma itu, misal dalam inisialisasi, fungsi penghitungan jarak, dan juga kriteria penghentian iterasi [1].

Berdasarkan latar belakang di atas, maka perlu dilakukan penelitian tentang mengimplementasikan algoritma K-Means kedalam sistem pendukung pengambilan keputusan yang dapat digunakan untuk membantu pengelompokan kemampuan siswa pada mata pelajaran Ujian Nasional.

II. LANDASAN TEORI

2.1 Algoritma K-Means

K-means merupakan metode pengklasteran secara *partitioning* yang memisahkan data ke dalam kelompok yang berbeda. Dengan *partitioning* secara iteratif, K-Means mampu meminimalkan rata-rata jarak setiap data ke klasternya. [4]. Dalam algoritma K-Means, setiap data harus termasuk ke *cluster* tertentu pada suatu tahapan proses, pada tahapan proses berikutnya dapat berpindah ke *cluster* yang lain [2]. Pada dasarnya penggunaan algoritma K-Means dalam melakukan proses *clustering* tergantung dari data yang ada dan konklusi yang ingin dicapai.

Untuk itu digunakan algoritma K-Means yang didalamnya memuat aturan sebagai berikut :

a. Jumlah *cluster* yang perlu di inputkan

b. Hanya memiliki atribut bertipe numerik

Algoritma K-Means pada awalnya mengambil sebagian dari banyaknya komponen dari populasi

untuk dijadikan pusat *cluster* awal. Pada step ini pusat *cluster* dipilih secara acak dari sekumpulan populasi data. Berikutnya K-Means menguji masing-masing komponen didalam populasi data dan menandai komponen tersebut ke salah satu pusat *cluster* yang telah di definisikan tergantung dari jarak minimum antar komponen dengan tiap-tiap pusat *cluster*. Posisi pusat *cluster* akan dihitung kembali sampai semua komponen data digolongkan kedalam tiap-tiap *cluster* dan terakhir akan terbentuk posisi *cluster* baru.

Algoritma K-Means pada dasarnya melakukan 2 proses yakni proses pendeteksian lokasi pusat *cluster* dan proses pencarian anggota dari tiap-tiap *cluster*. Proses *clustering* dimulai dengan mengidentifikasi data yang akan dikluster, X_{ij} ($i=1, \dots, n$; $j=1, \dots, m$) dengan n adalah jumlah data yang akan dikluster dan m adalah jumlah variabel. Pada awal iterasi, pusat setiap kluster ditetapkan secara bebas (sembarang), C_{kj} ($k=1, \dots, k$; $j=1, \dots, m$). Kemudian dihitung jarak antara setiap data dengan setiap pusat *cluster*. Untuk melakukan penghitungan jarak data ke- i (x_i) pada pusat *cluster* ke- k (c_k), diberi nama (d_{ik}), dapat digunakan formula Euclidean. Suatu data akan menjadi anggota dari *cluster* ke- k apabila jarak data tersebut ke pusat *cluster* ke- k bernilai paling kecil jika dibandingkan dengan jarak ke pusat *cluster* lain [6].

Proses dasar algoritma K-Means [8] :

1. Tentukan k sebagai jumlah *cluster* yang ingin dibentuk. Tetapkan pusat *cluster*.
2. Hitung jarak setiap data ke pusat *cluster* menggunakan persamaan Euclidean.

$$d_{ik} = \sqrt{\sum_j^m (C_{ij} - C_{kj})^2} \quad (1)$$

3. Kelompokkan data ke dalam *cluster* yang dengan jarak yang paling pendek menggunakan persamaan.

$$\text{Min} \sum_{k=1}^k d_{ik} = \sqrt{\sum_j^m (C_{ij} - C_{kj})^2} \quad (2)$$

4. Hitung pusat *cluster* yang baru menggunakan persamaan

$$C_{kj} = \frac{\sum_{i=1}^p x_{ij}}{p} \quad (3)$$

Dimana :

$$x_{ij} \in \text{kluster ke - } k$$

P = banyaknya anggota *cluster* ke k

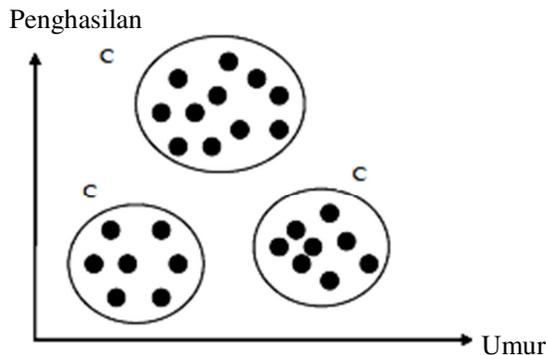
- Ulangi langkah 2 sampai dengan 4 hingga sudah tidak ada lagi data yang berpindah ke cluster yang lain.

2.2 Pengklasteran (Clustering)

Clustering adalah suatu alat untuk analisa data, yang memecahkan permasalahan pengelompokan. Obyeknya ialah untuk kasus pendistribusian (orang-orang, objek, peristiwa dan lainnya) ke dalam kelompok, sedemikian hingga derajat tingkat keterhubungan antar anggota cluster yang sama adalah kuat dan lemah antar anggota dari cluster yang berbeda. Dengan Cara ini masing-masing cluster menguraikan, dalam kaitan dengan kumpulan atau koleksi data, class dimana milik anggota-anggotanya. Cluster disebut juga data item dikelompokkan menurut pilihan konsumen atau hubungan logis.

Analisis kluster [3] ialah metode yang dipakai untuk membagi rangkaian data menjadi beberapa grup berdasarkan kesamaan-kesamaan yang telah ditentukan sebelumnya. Jadi menurut Gorunescu (2011) secara umum dapat dikatakan bahwa :

- Data dalam satu cluster memiliki tingkat kesamaan yang tinggi, dan
 - Data dalam cluster yang berbeda memiliki tingkat kesamaan yang rendah
- Untuk lebih jelasnya perhatikan gambar berikut :



Gambar 1. Contoh Clustering (Larose, 2005)

Dari gambar 1, kita misalkan sebagai kumpulan data konsumen sederhana yang mengandung dua atribut yaitu umur dan penghasilan. Berdasarkan atribut tersebut kemudian terbagi menjadi tiga kelompok (*cluster*) yaitu *C1* yang terdiri konsumen usia muda dengan penghasilan rendah. *C2* terdiri dari konsumen usia muda dan tua dengan penghasilan tinggi. *C3* terdiri dari konsumen usia tua dengan penghasilan relatif rendah.

III. PERHITUNGAN MANUAL DARI FUNGSI ALGORITMA K-MEANS

Percobaan dilakukan dengan menggunakan parameter-parameter berikut:

Jumlah Cluster : 4

Jumlah Data : 12

Jumlah Atribut : 4 (IPA, MTK, B. Indonesia, dan B. Inggris)

| NIS | NAMA | RATA-RATA | | | |
|-------|----------------|-----------|------|------|------|
| | | BIND | BING | IPA | MTK |
| 10943 | ARIS APRISALY | 8.1 | 7.3 | 7.76 | 7.42 |
| 10944 | ARMANDA D | 8.17 | 7.33 | 7.64 | 7.36 |
| 10945 | AYU INDAH L | 8.5 | 7.45 | 7.88 | 7.59 |
| 10904 | AZIZAH LARAS | 8.65 | 8.42 | 8.18 | 8.13 |
| 10835 | CINDY AULIA | 8.08 | 7.37 | 8.02 | 7.54 |
| 10800 | DENY OKTA | 7.27 | 8 | 8 | 8 |
| 11012 | DHEA AMEILYA | 8.23 | 7.33 | 7.28 | 7.69 |
| 11092 | ESA SALSZA Z | 8.52 | 7.82 | 8.22 | 7.81 |
| 10913 | FARRAS JIHAN A | 8.43 | 7.85 | 7.9 | 7.87 |
| 10876 | FAUZAN HANIF | 8.15 | 7.58 | 7.88 | 7.47 |
| 11020 | GAYUSMAN P | 8.12 | 7.42 | 7.55 | 7.35 |

Iterasi ke-1

- Penentuan pusat awal *cluster*.

Untuk penentuan awal diasumsikan dengan menggunakan nilai terendah pada masing-masing mata pelajaran :

| | | | | |
|---------|------|------|-------|------|
| DATA 6 | 7.27 | 8 | 8 | 8 |
| DATA 1 | 8.1 | 7.3 | 7.758 | 7.42 |
| DATA 7 | 8.23 | 7.33 | 7.28 | 7.69 |
| DATA 12 | 8.3 | 7.42 | 7.658 | 7.34 |

- Perhitungan jarak pusat *cluster*

Untuk mengukur jarak antara data dengan pusat *cluster* digunakan *Euclidian distance*, kemudian akan didapatkan matriks jarak sebagai berikut :

Rumus *Euclidian distance* :

$$d_{ik} = \sqrt{\sum_j^m (C_{ij} - C_{kj})^2}$$

C_{ij} : Pusat Cluster

C_{kj} : Data

Sebagai contoh, perhitungan jarak dari data ke-1 terhadap pusat *cluster* adalah :

$$C1 = \sqrt{(8.1 - 7.27)^2 + (7.3 - 8)^2 + (7.758 - 8)^2 + (7.42 - 8)^2}$$

$$C2 = \sqrt{(8.1 - 8.1)^2 + (7.3 - 7.3)^2 + (7.758 - 7.758)^2 + (7.42 - 7.42)^2}$$

$$C3 = \sqrt{(8.1 - 8.23)^2 + (7.3 - 7.33)^2 + (7.758 - 7.28)^2 + (7.42 - 7.69)^2}$$

$$C4 = \sqrt{(8.1 - 8.3)^2 + (7.3 - 7.42)^2 + (7.758 - 7.658)^2 + (7.42 - 7.34)^2}$$

Dan seterusnya dilanjutkan untuk data ke 2, ... N
Kemudian akan didapatkan matriks jarak sebagai berikut :

D1 =

Iterasi ke-1

| | C1 | C2 | C3 | C4 |
|---------------------|------|------|------|------|
| ARIS APRISALY | 1.26 | 0 | 0.57 | 0.26 |
| ARMANDA DWICAHYA | 1.34 | 0.15 | 0.5 | 0.16 |
| AYU INDAH LESTARI | 1.41 | 0.48 | 0.67 | 0.39 |
| AZIZAH LARASWATI | 1.46 | 1.5 | 1.53 | 1.42 |
| CINDY AULIA | 1.13 | 0.3 | 0.77 | 0.47 |
| DENY OKTAVIANTO | 0 | 1.26 | 1.41 | 1.4 |
| DHEA AMEILYA | 1.41 | 0.57 | 0 | 0.53 |
| ESA SALSA ZENITA | 1.29 | 0.9 | 1.1 | 0.86 |
| FARRAS JIHAN AFIFAH | 1.18 | 0.8 | 0.85 | 0.73 |
| FAUZAN HANIF | 1.12 | 0.31 | 0.69 | 0.34 |
| GAYUSMAN PUTRANTA | 1.3 | 0.25 | 0.46 | 0.21 |
| GOVINDA PANGESTU | 1.4 | 0.26 | 0.53 | 0 |

Setiap kolom pada matrik menunjukkan nilai jarak data terhadap pusat *cluster*. Baris pertama pada matrik menunjukkan nilai jarak data terhadap titik pusat *cluster* pertama, baris kedua pada matrik menunjukkan nilai jarak data terhadap titik pusat *cluster* kedua dan seterusnya.

3. Pengelompokan data

Jarak hasil perhitungan akan dilakukan perbandingan dan dipilih jarak terdekat antara data dengan pusat cluster, jarak ini menunjukkan bahwa data tersebut berada dalam satu kelompok dengan pusat cluster terdekat. Berikut ini akan ditampilkan data matriks pengelompokan group, nilai 1 berarti data tersebut berada dalam group.

G1=

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|----|---|---|---|---|---|---|---|---|---|----|----|----|
| C1 | | | | | | 1 | | | | | | |
| C2 | 1 | 1 | | | 1 | | | | 1 | 1 | | |
| C3 | | | | | | | 1 | | | | | |
| C4 | | | 1 | 1 | | | | 1 | | | 1 | 1 |

Keterangan :

- Jika nilai D1 lebih dekat jaraknya dari nilai hasil perhitungan data *cluster* ke-1 (C1) atau data *cluster* ke-2 (C2) atau data *cluster* ke-3 (C3) atau data *cluster* ke-4 (C4) maka G1 bernilai 0 dan tidak termasuk grup atau kelompok *cluster* baru.
- Jika nilai D1 lebih jauh jaraknya dari nilai hasil perhitungan data *cluster* ke-1 (C1) atau data *cluster* ke-2 (C2) atau data *cluster* ke-3 (C3) atau data *cluster* ke-4 (C4) maka G1 bernilai 1 dan termasuk grup atau kelompok *cluster* baru.
- Berdasarkan matriks data perhitungan, didapat :

C1 data 6

C3 data 7

C2 data 1,2,5,9,10

C4 data 3,4,8,11,12

4. Penentuan pusat cluster baru

Setelah diketahui anggota tiap-tiap cluster kemudian pusat cluster baru dihitung berdasarkan data anggota tiap-tiap cluster sesuai dengan rumus pusat anggota cluster. Sehingga didapatkan perhitungan sebagai berikut :

Karena C1 memiliki 1 anggota yaitu data 6 maka perhitungan *cluster* baru menjadi :

$$C1 = \left(\frac{7.27}{1}, \frac{8}{1} \right)$$

Lakukan hal yang sama pada C2, C3, dan C4 .
Nilai *cluster* yang baru adalah :

Cluster baru untuk iterasi ke-2

| C1 | 7.27 | 8 | 8 | 8 |
|----|------|------|------|------|
| C2 | 8.19 | 7.49 | 7.84 | 7.53 |
| C3 | 8.23 | 7.33 | 7.28 | 7.69 |
| C4 | 8.42 | 7.7 | 7.9 | 7.64 |

Iterasi ke-2

5. Ulangi langkah ke 2 (kedua) hingga posisi data tidak mengalami perubahan.

D² =

| | C1 | C2 | C3 | C4 |
|----|-------|-------|-------|------|
| 1 | 1.256 | 0.248 | 0.567 | 0.58 |
| 2 | 1.337 | 0.303 | 0.495 | 0.59 |
| 3 | 1.413 | 0.324 | 0.672 | 0.27 |
| 4 | 1.459 | 1.249 | 1.536 | 0.94 |
| 5 | 1.128 | 0.239 | 0.767 | 0.5 |
| 6 | 0 | 1.162 | 1.409 | 1.24 |
| 7 | 1.409 | 0.603 | 0.005 | 0.74 |
| 8 | 1.293 | 0.662 | 1.1 | 0.39 |
| 9 | 1.185 | 0.557 | 0.853 | 0.27 |
| 10 | 1.117 | 0.127 | 0.689 | 0.34 |
| 11 | 1.297 | 0.354 | 0.457 | 0.61 |
| 12 | 1.397 | 0.293 | 0.526 | 0.49 |

G²=

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|----|---|---|---|---|---|---|---|---|---|----|----|----|
| C1 | | | | | | 1 | | | | | | |
| C2 | 1 | 1 | | | 1 | | | | | 1 | 1 | 1 |
| C3 | | | | | | | 1 | | | | | |
| C4 | | | 1 | 1 | | | | 1 | 1 | | | |

C1 data 6

C3 data 7

C2 data 1,2,5,10,11,12

C4 data 3,4,8,9

Iterasi akan terus dilakukan hingga nilai G pada tiap iterasi memiliki nilai yang sama. Pada saat nilai G yang sama telah terjadi, maka clustering telah

mencapai stabil dan konvergen. Dan pada proses penghitungan manual ini, nilai G yang konvergen terjadi di iterasi ke-4, pada iterasi ke-4, nilai $G^3=G^4$. G^3 dan G^4

| | | | | | | | | | | | | |
|----|---|---|---|---|---|---|---|---|---|----|----|----|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| C1 | | | | | | 1 | | | | | | |
| C2 | 1 | 1 | 1 | | 1 | | | | | 1 | 1 | 1 |
| C3 | | | | | | | 1 | | | | | |
| C4 | | | | 1 | | | | 1 | 1 | | | |

C1 data 6
C2 data 1,2,3,5,10,11,12
C3 data 7
C4 data 3,4,8,9

IV. IMPLEMENTASI ALGORITMA K-MEANS PADA APLIKASI

4.1. Input

Input data dilakukan oleh admin dan user yang diberi wewenang untuk melakukan input data. Data yang harus diinput untuk menggunakan aplikasi adalah data siswa dan data nilai semester satu hingga lima.

Gambar 2. Halaman Input Data Siswa

Gambar 3. Halaman Input Nilai

4.2. Output

Output yang diharapkan dari aplikasi ini adalah terbentuknya kelompok pengayaan materi berdasarkan kemampuan siswa atau kebutuhan materi yang perlu diajarkan pada siswa. Output aplikasi disajikan dalam dua bentuk yaitu tampilan pada layar dan report yang dapat dicetak di printer.

Gambar 4. Hasil Pengelompokan Kemampuan

Gambar 4 menampilkan hasil iterasi algoritma K-Means untuk mengelompokkan siswa berdasarkan kemampuan siswa terhadap materi yang akan diujikan pada ujian nasional. Jumlah siswa yang masuk kedalam kelompok setiap materi ditunjukkan pada gambar berikutnya.

Gambar 5. Halaman Jumlah Hasil Pengelompokan

Gambar 5 menunjukkan empat kelompok materi pengayaan yaitu Bahasa Indonesia terdapat 64 siswa, Bahasa Inggris sebanyak 93 siswa, IPA terdiri 46 siswa dan Matematika ada 72 siswa. Berdasarkan hasil tersebut, maka bisa jadi ada siswa yang harus mengikuti seluruh materi pengayaan, ada juga yang hanya mengikuti satu atau dua materi saja. Output kemampuan siswa secara keseluruhan dapat dicetak, tampilan sebelum dicetak adalah sebagai berikut:

Gambar 6 Halaman Print Hasil Pengelompokan Kemampuan

Hasil pengelompokan ini dapat selain dimanfaatkan untuk menentukan kelompok pengayaan materi, dapat juga digunakan untuk mengukur progres siswa

terhadap materi yang sudah dikuasai. Nilai semester 6 yang berasal dari nilai *Try Out UN* yang diadakan pihak sekolah bersama sebuah bimbingan belajar diinput dan diolah hingga menghasilkan kelompok yang baru. Hasil pengelompokan ini dibandingkan dengan hasil pengelompokan sebelumnya hingga terlihat perkembangan kemampuan siswa berdasar jumlah kelompok materi yang harus diikuti.

V. PENUTUP

5.1 Kesimpulan

1. Algoritma K-Means dapat diimplementasikan untuk membantu pengelompokan kemampuan siswa terhadap mata pelajaran Ujian Nasional.
2. Aplikasi ini juga dapat melihat perkembangan kemampuan siswa setelah mengikuti pengayaan materi.

5.2 Saran

Beberapa saran guna pengembangan aplikasi ini agar menjadi lebih baik lagi, yaitu :

1. Jumlah sampel data harus diperbanyak dan diukur waktu pengolahannya.
2. Perlu dilakukan uji coba dengan algoritma *clustering* lain seperti K-Means ++.

DAFTAR PUSTAKA

- [1] Celebi, M. E., Kingravi, H. A., & Vela, P. A. 2013. A comparative study of efficient initialization methods for the k-means clustering algorithm. *Expert Systems with Applications*
- [2] Ghosh, Soumi dan Dubey, Sanjay Kumar. 2013. *Comparative Analysis of K-Means and Fuzzy C-Means Algorithm*. India
- [3] Han, Jiawei., Kamber, Micheline., Pei, Jian. 2011. *Data Mining Concept and Techniques Third*

Edition. San Francisco: Morgan Kaufmann Publishers

- [4] MacQueen, J. B. 1967. Some Methods for classification and Analysis of Multivariate Observations, *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley : University of California Press
- [5] Marini, Vivi. 2013. *Efektivitas Bimbingan Belajar Menghadapi Ujian Nasional Pelajaran Sosiologi SMA Muhammadiyah 2 Pontianak*. Jurnal. Pontianak: Universitas Tanjungpura
- [6] Panda, Sandeep., Sahu, Sanat., Jena, Pradeep., & Chattopadhyay, Subhagata. 2013 *Comparing Fuzzy C-Means and K-Means Clustering Techniques: A Comprehensive Study*. India
- [7] Rokach, L. 2010. A survey of Clustering Algorithms. In O. Maimon & L. Rokach (Eds.), *Data Mining and Knowledge Discovery Handbook (2nd ed.)*. Springer Science+Business Media, LLC.
- [8] Aryan, Peb Rusyono. 2010. *Algoritma K-means Clustering*. <http://pebbie.wordpress.com/2008/11/13/algoritma-kmeansclustering/> Diakses pada tanggal 18 Agustus 2014 pukul 13.47 WIB
- [9] indopos.co.id/2014/04/dki-siap-hadapi-ujian-nasional.html diakses pada tanggal 12 November 2014 pukul 15.18 WIB