

Mengelompokkan Siswa Berprestasi Akademik dengan Menggunakan Metode *K Means* Kelas VII MTs Hidayatul Mubtadi'in Pancoran Kadur

Nur Jannah¹, Tony Yulianto²

^{1,2}Jurusan Matematika, FMIPA Universitas Islam Madura (UIM)
Jl. Bettet No. 04, Pamekasan, Madura 60111 Indonesia
Email: nurjannah@gmail.com

ABSTRAK

Pengelompokan atau lazim di kenal dengan *grouping* didasarkan atas pandangan bahwa disamping peserta didik tersebut mempunyai kesamaan, dan juga mempunyai perbedaan. Pengelompokan bukan hanya melihat dari hasil akhir rapor semester i dan semester ii akan tetapi juga dilihat dari nilai kognitif, psikomotorik dan afektif selama proses belajar berlangsung. Proses pengelompokkan siswa berprestasi untuk mempermudah dalam proses belajar mengajar. Melaksanakan pengelompokan siswa berprestasi masih belum maksimal karena kemampuan siswa tersebut tidak hanya diukur dari rapor terakhir saja, melainkan nilai kognitif, psikomotorik dan afektifnya. Berikut nilai tes terakhir untuk menambah keakuratan data. Pengelompokan dari 30 siswa dengan kriteria pembobot pintar, sedang, dan tidak pintar. Untuk itu diperlukan solusi yang dapat mengatasi kesulitan tersebut. Metode *clustering* dengan menggunakan algoritma *k means* diimplementasikan dalam aplikasi ini., juga dapat digunakan untuk memantau perkembangan kemampuan setelah mengikuti belajar siswa. Dari hasil dan pembahasan dapat diperoleh hasil dari validasinya data valid ada 33.3333% sedangkan data tidak valid ada 66.6667%.

Kata kunci: *Ujian Semester I Dan Semester II, Clustering, Algoritma K-Means*

1. PENDAHULUAN

Pendidikan merupakan usaha sadar dan terencana untuk mewujudkan suasana belajar dan proses pembelajaran, agar peserta didik secara aktif mengembangkan potensi dirinya untuk memiliki kekuatan spiritual keagamaan, pengendalian diri, kecerdasan, akhlak mulia, serta keterampilan yang diperlukannya masyarakat, bangsa dan negara. Dari pengertian ini, pendidikan adalah modal utama realita kehidupan bermasyarakat, berbangsa dan bernegara (Arifin dan Fariyah: 2012).

Prestasi belajar banyak diartikan sebagai seberapa jauh hasil yang telah dicapai siswa dalam penguasaan tugas-tugas atau materi pelajaran yang diterima dalam jangka waktu tertentu. Bagi siswa prestasi merupakan sesuatu yang amat penting lebih khusus lagi dalam hal prestasi belajar, karena nilai yang dicapai dalam proses belajar adalah prestasi yang dapat dilihat secara nyata. Untuk mendapatkan prestasi yang baik maka di perlukan pengelompokan peserta didik.

Pengelompokan atau lazim di kenal dengan *grouping* didasarkan atas pandangan bahwa disamping peserta didik tersebut mempunyai kesamaan, juga mempunyai perbedaan. Kesamaan-kesamaan yang ada pada peserta didik melahirkan pemikiran penempatan pada kelompok yang sama, sementara perbedaan-perbedaan yang ada pada

peserta didik melahirkan pemikiran pengelompokan mereka pada kelompok yang berbeda (Azhar, 2012).

Perkembangan teknologi informasi yang pesat sekarang ini membawa banyak perubahan dalam pola kehidupan masyarakat. Sekarang teknologi informasi juga menunjang dalam penyelesaian pekerjaan secara cepat, efektif dan efisien. Misalnya saja pada pengelompokan data atau disebut juga dengan *clustering*.

Data *clustering* merupakan salah satu metode data *mining* yang bersifat tanpa arahan. *K-Means* merupakan metode data *clustering* non hirarki yang berusaha mempartisi data yang ada dalam bentuk satu atau lebih *cluster*/kelompok. Metode ini mempartisi data dalam *cluster*/kelompok sehingga data yang memiliki karakteristik yang sama dikelompokkan dalam satu *cluster* yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan dalam kelompok yang lain. Metode *K-Means* ini secara umum dilakukan dengan algoritma dasar, (Ahmadi dan Hartati, 2010).

2. TINJAUAN PUSTAKA

2.1 Penelitian Sebelumnya

Penelitian tentang *K Means* yang pernah dilakukan antara lain "Penerapan algoritma *K Means* pada kompresi adaptif citra medis MRI". (Angga, 2015)

Dalam sebuah penelitian sebelumnya yang juga disusun oleh Narwati [2] tentang “Pen gelompokan mahasiswa menggunakan algoritma *K Means*” didalamnya menjelaskan bahwa peneliti ingin menseksi masuknya mahasiswa dalam sebuah perguruan tinggi umumnya dengan memberikan soal-soal test yang harus mereka kerjakan, untuk mengetahui kemampuan dan pengetahuan mereka (Narwati, 2006).

2.2 Tinjauan Tentang Prestasi Belajar

Prestasi merupakan hasil akhir dari pekerjaan yang telah dilakukan. Prestasi tidak akan diperoleh sebelum seseorang melakukan kegiatan. Bagi siswa prestasi merupakan sesuatu yang amat penting lebih khusus lagi dalam hal prestasi belajar, karena nilai yang dicapai dalam proses belajar adalah prestasi yang dapat dilihat secara nyata (Setiawan, 2009).

2.3 Data Mining

Data mining adalah suatu istilah yang digunakan untuk menemukan pengetahuan yang tersembunyi di dalam database. Data mining merupakan proses semi otomatis yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi pengetahuan potensial dan berguna yang bermanfaat yang tersimpan di dalam database besar (Larose, 2006).

2.4 Pengklasteran (Clustering)

Clustering adalah suatu alat untuk analisa data yang memecahkan permasalahan pengelompokan. Obyeknya merupakan 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 (Larose, 2006).

2.5 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 *klaster*nya. Algoritma *K Means* (Agustin, Fitria, & Hanifah, 2009).

Tentukan *k* sebagai jumlah *cluster* yang ingin dibentuk. Tetapkan pusat *cluster* awal.

Hitung jarak setiap data (*x*) ke pusat *cluster* (*c*) menggunakan persamaan *Euclidean*

$$d_{ik} = \sqrt{\sum_j^m (x_{ij} - c_{kj})^2} \quad (2.1)$$

Kelompokkan data dari persamaan (2.1) ke dalam *cluster* yang dengan jarak yang paling pendek menggunakan persamaan.

$$S_i = \min_i d_{ik} \quad (2.2)$$

Update pusat *cluster* yang baru dari rata-rata *cluster* yang bersangkutan menggunakan persamaan.

$$C_{kj} = \frac{\sum_{i=1}^n x_{ij}}{p} \quad (2.3)$$

Dengan:

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

$$P = \text{banyaknya anggota cluster ke } k.$$

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

$$C_{\text{new}} - C_{\text{old}} < \varepsilon$$

Dengan:

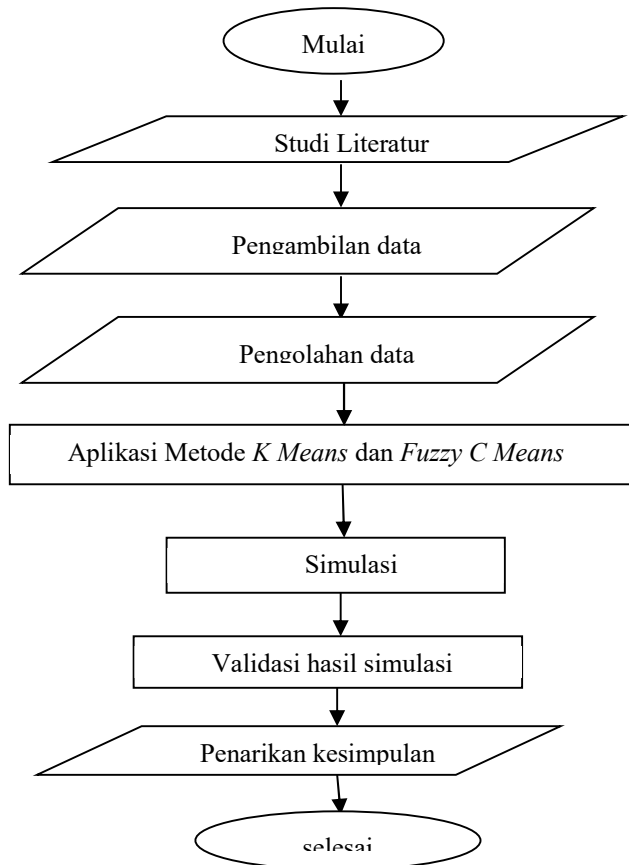
$$C_{\text{new}} = \text{pusat cluster baru}$$

$$C_{\text{old}} = \text{pusat cluster lama}$$

$$\varepsilon = 10^{-5}$$

3. METODE PENELITIAN

Pada subbab ini akan dijelaskan tentang metode yang digunakan dalam penelitian ini disertai dengan pustaka yang mendasari teori dalam penelitian ini, seperti penelitian sebelumnya, pengertian pendidikan, pengelompokan, dan metode *K Means*. Adapun untuk langkah-langkah dalam penelitian ini dapat dilihat dalam Gambar 1.



Gambar 1. Flowchart Penelitian

4. ANALISIS HASIL DAN PEMBAHASAN

Dengan menggunakan metode *K Means* pada pengelompokan siswa berprestasi akademik untuk

kelas VII maka sebagai berikut:

A. Metode K Means untuk kelas VII A

1. Muzayyanah

$$\begin{aligned} i=1 \quad k=1 \quad d_{11} &= \sqrt{\sum_{j=1}^3 (x_{1j} - c_{1j})^2} \\ &= \sqrt{(x_{11} - c_{11})^2 + (x_{12} - c_{12})^2 + (x_{13} - c_{13})^2} \\ &= \sqrt{(78.46 - 0.8)^2 + (83.5 - 0.15)^2 + (75 - 0.05)^2} \\ &= \sqrt{6031.07 + 6947.22 + 5617.50} \\ &= \sqrt{18595.79} \\ &= 136.3664 \end{aligned}$$

$$\begin{aligned} i=1 \quad k=2 \quad d_{12} &= \sqrt{\sum_{j=1}^3 (x_{1j} - c_{2j})^2} \\ &= \sqrt{(x_{11} - c_{21})^2 + (x_{12} - c_{22})^2 + (x_{13} - c_{23})^2} \\ &= \sqrt{(78.46 - 0.5)^2 + (83.5 - 0.3)^2 + (75 - 0.2)^2} \\ &= \sqrt{6077.76 + 6922.24 + 5595.04} \\ &= \sqrt{18595.04} \\ &= 136.3636 \end{aligned}$$

$$\begin{aligned} i=1 \quad k=3 \quad d_{13} &= \sqrt{\sum_{j=1}^3 (x_{1j} - c_{3j})^2} \\ &= \sqrt{(x_{11} - c_{31})^2 + (x_{12} - c_{32})^2 + (x_{13} - c_{33})^2} \\ &= \sqrt{(78.46 - 0.25)^2 + (83.5 - 0.25)^2 + (75 - 0.5)^2} \\ &= \sqrt{6116.80 + 6930.56 + 5550.25} \\ &= \sqrt{18597.61} \\ &= 136.3731 \end{aligned}$$

2. Nailatus Sarofah

$$\begin{aligned} i=2 \quad k=1 \quad d_{21} &= \sqrt{\sum_{j=1}^3 (x_{2j} - c_{1j})^2} \\ &= \sqrt{(x_{21} - c_{11})^2 + (x_{22} - c_{12})^2 + (x_{23} - c_{13})^2} \\ &= \sqrt{(77.21 - 0.8)^2 + (80 - 0.15)^2 + (82.25 - 0.05)^2} \\ &= \sqrt{5838.48 + 6376.02 + 6756.84} \\ &= \sqrt{18971.34} \\ &= 137.7389 \end{aligned}$$

$$\begin{aligned} i=2 \quad k=2 \quad d_{22} &= \sqrt{\sum_{j=1}^3 (x_{2j} - c_{2j})^2} \\ &= \sqrt{(x_{21} - c_{21})^2 + (x_{22} - c_{22})^2 + (x_{23} - c_{23})^2} \\ &= \sqrt{(77.21 - 0.5)^2 + (80 - 0.3)^2 + (82.25 - 0.2)^2} \\ &= \sqrt{5884.42 + 6352.09 + 6732.20} \\ &= \sqrt{18968.71} \\ &= 137.7294 \end{aligned}$$

$$\begin{aligned} i=2 \quad k=3 \quad d_{23} &= \sqrt{\sum_{j=1}^3 (x_{2j} - c_{3j})^2} \\ &= \sqrt{(x_{21} - c_{31})^2 + (x_{22} - c_{32})^2 + (x_{23} - c_{33})^2} \\ &= \sqrt{(77.21 - 0.25)^2 + (80 - 0.25)^2 + (82.25 - 0.5)^2} \\ &= \sqrt{65922.84 + 6360.06 + 6683.06} \\ &= \sqrt{18965.96} \\ &= 137.7194 \end{aligned}$$

3. Paiqotul Jannah

$$\begin{aligned} i=3 \quad k=1 \quad d_{31} &= \sqrt{\sum_{j=1}^3 (x_{3j} - c_{1j})^2} \\ &= \sqrt{(x_{31} - c_{11})^2 + (x_{32} - c_{12})^2 + (x_{33} - c_{13})^2} \end{aligned}$$

$$\begin{aligned} &= \sqrt{(77.29 - 0.8)^2 + (85.5 - 0.15)^2 + (73.5 - 0.05)^2} \\ &= \sqrt{5850.72 + 7284.62 + 5394.90} \\ &= \sqrt{18530.24} \\ &= 136.1234 \end{aligned}$$

$$\begin{aligned} i=3 \quad k=2 \quad d_{32} &= \sqrt{\sum_{j=1}^3 (x_{3j} - c_{2j})^2} \\ &= \sqrt{(x_{31} - c_{21})^2 + (x_{32} - c_{22})^2 + (x_{33} - c_{23})^2} \\ &= \sqrt{(77.29 - 0.5)^2 + (85.5 - 0.3)^2 + (73.5 - 0.2)^2} \\ &= \sqrt{5896.70 + 7259.04 + 5372.89} \\ &= \sqrt{18528.63} = 136.1175 \end{aligned}$$

$$\begin{aligned} i=3 \quad k=3 \quad d_{33} &= \sqrt{\sum_{j=1}^3 (x_{3j} - c_{3j})^2} \\ &= \sqrt{(x_{31} - c_{31})^2 + (x_{32} - c_{32})^2 + (x_{33} - c_{33})^2} \\ &= \sqrt{(77.29 - 0.25)^2 + (85.5 - 0.25)^2 + (73.5 - 0.5)^2} \\ &= \sqrt{5935.16 + 7267.56 + 5329} \\ &= \sqrt{18531.72} \\ &= 136.1288 \end{aligned}$$

Update pusat cluster dengan persamaan

$$C_{kj} = \frac{\sum_{i=1}^n x_{ij}}{p}$$

Ambil k = 1 dan j = 1 diperoleh $C_{11} = \frac{\sum_{i=1}^n x_{ij}}{p}$

$$\begin{aligned} &= \frac{x_{11} + x_{21} + x_{31} + x_{41} + x_{51}}{p} \\ &= \frac{78.46 + 77.21 + 77.29 + 76.96 + 78}{1} \\ &= \frac{387.92}{1} \\ &= 387.92 \end{aligned}$$

$$\begin{aligned} C_{12} &= \frac{\sum_{i=1}^n x_{ij}}{p} \\ &= \frac{x_{12} + x_{22} + x_{32} + x_{42} + x_{52}}{p} \\ &= \frac{83.5 + 80 + 85.5 + 83 + 80}{1} \\ &= \frac{412}{1} \\ &= 412 \end{aligned}$$

$$\begin{aligned} C_{13} &= \frac{\sum_{i=1}^n x_{ij}}{p} \\ &= \frac{x_{13} + x_{23} + x_{33} + x_{43} + x_{53}}{p} \\ &= \frac{75 + 82.25 + 73.5 + 75 + 81.5}{1} \\ &= \frac{387.25}{1} \\ &= 387.25 \end{aligned}$$

Amil k = 2 dan j = 1 diperoleh $C_{21} = \frac{\sum_{i=1}^n x_{ij}}{p}$

$$\begin{aligned} &= \frac{x_{11} + x_{21} + x_{31} + x_{41} + x_{51}}{p} \\ &= \frac{78.46 + 77.21 + 77.29 + 76.96 + 78}{5} \\ &= \frac{387.92}{5} \\ &= 77.58 \end{aligned}$$

$$\begin{aligned} C_{22} &= \frac{\sum_{i=1}^n x_{ij}}{p} \\ &= \frac{x_{12} + x_{22} + x_{32} + x_{42} + x_{52}}{p} \end{aligned}$$

$$= \frac{83.5+80+85.5+83+80}{5}$$

$$= \frac{412}{5}$$

$$= 82.4$$

$$C_{23} = \frac{\sum_{i=1}^n x_{ij}}{p}$$

$$= \frac{x_{13}+x_{23}+x_{33}+x_{43}+x_{53}}{p}$$

$$= \frac{75+82.25+73.5+75+81.5}{5}$$

$$= \frac{387.25}{5}$$

$$= 77.45$$

Ambil k = 3 dan j = 1 diperoleh $C_{31} = \frac{\sum_{i=1}^n x_{ij}}{p}$

$$= \frac{x_{11}+x_{21}+x_{31}+x_{41}+x_{51}}{p}$$

$$= \frac{78.46+77.21+77.29+76.96+78}{24}$$

$$= \frac{387.92}{24}$$

$$= 16.16$$

$$C_{32} = \frac{\sum_{i=1}^n x_{ij}}{p}$$

$$= \frac{x_{12}+x_{22}+x_{32}+x_{42}+x_{52}}{p}$$

$$= \frac{83.5+80+85.5+83+80}{24}$$

$$= \frac{412}{24}$$

$$= 17.16$$

$$C_{33} = \frac{\sum_{i=1}^n x_{ij}}{p}$$

$$= \frac{x_{13}+x_{23}+x_{33}+x_{43}+x_{53}}{p}$$

$$= \frac{75+82.25+73.5+75+81.5}{24}$$

$$= \frac{387.25}{24}$$

$$= 16.13$$

$$\text{Error} = |C_{11 \text{ new}} - C_{11 \text{ old}}| + |C_{12 \text{ new}} - C_{12 \text{ old}}| +$$

$$|C_{13 \text{ new}} - C_{13 \text{ old}}| + |C_{21 \text{ new}} - C_{21 \text{ old}}| +$$

$$|C_{22 \text{ new}} - C_{22 \text{ old}}| + |C_{23 \text{ new}} - C_{23 \text{ old}}| +$$

$$|C_{31 \text{ new}} - C_{31 \text{ old}}| + |C_{32 \text{ new}} - C_{32 \text{ old}}| +$$

$$|C_{33 \text{ new}} - C_{33 \text{ old}}|$$

$$= |0.8 - 387.92| + |0.15 - 412| + |0.05 - 387.25| +$$

$$|0.5 - 77.58| + |0.3 - 82.4| + |0.2 - 77.45| +$$

$$|0.25 - 16.16| + |0.25 - 16.17| + |0.5 - 16.13|$$

$$= 1470.15$$

Untuk pengerjaan iterasi selanjutnya sama dengan sebelumnya.

Hasil dari validasi hasil simulasi dapat dilihat pada Tabel 1:

Tabel 1 Data Hasil Simulasi

No	Nama	KELOMPOK
1	Muzayyanah	Pintar
2	Lailatis Sarofah	Sedang
3	Paiqotul Jannah	Sedang
4	Anas	Sedang
5	Fathur Rosi B	Sedang
6	Halimatus Zahroh	Pintar
7	Sitti Mutmainnah	Pintar
8	Wahid	Tidak Pintar
9	Sirojuddin	Pintar
10	Nur Holis	Tidak Pintar
11	Sulastri	Pintar
12	Hudaifah	Tidak Pintar
13	Zainuddin	Sedang
14	Saidi	Sedang
15	Atiatul Hasanah	Sedang
16	Dewi Ulfami	Sedang
17	Aisyah	Sedang
18	Rizkiyatul Hasanah	Pintar
19	Fatimatus Zahroh	Pintar
20	Makkiyah	Pintar
21	Halimatus Saadah	Pintar
22	Fajarullah	Pintar
23	Dandi Yanto	Tidak Pintar
24	Fathur Rosi A	Tidak Pintar
25	Hendri	Tidak Pintar
26	Rofiqoh	Tidak Pintar
27	Uswatun Hasanah	Tidak Pintar
28	Mutmainah	Sedang
29	Irofatul Imamah	Sedang
30	Flrdausiy Al Aniq	Sedang

5. KESIMPULAN

Dari hasil dan pembahasan diperoleh hasil dari validasinya data valid ada 33.3333% sedangkan data tidak valid ada 66.6667%.

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

1. Percobaan dengan menggunakan algoritma dengan pembelajaran sebelumnya (supervised learning)
2. Perlu dilakukan uji coba dengan algoritma clustering lain seperti *K Means* ++

DAFTAR PUSTAKA

Arifin, dan Fariyah. (2012). *Pengaruh Strategi Pembelajaran CTL Terhadap Motivasi Belajar Siswa Kelas VIII MTs*. Pamekasan: STAI AL-Khairot.

Ahmadi, A., dan Hartati, S. (2010). *Penerapan Fuzzy C-Means dalam Sistem Pendukung Keputusan untuk Penentuan Penerima Bantuan*

Langsung Masyarakat (BLM) PNPM MPd (Studi Kasus PNPM-MPd Kec. Ngadirojo Kab. Pacitan). Yogyakarta: Universitas Gadjah Mada Yogyakarta.

Azhar. (2012). *Definisi, Pengertian Dan Faktor-faktor Yang Mempengaruhi Prestasi Belajar*. Surabaya: Jakarta.

Larose. (2006). *Penerapan Data Mining Untuk Menentukan Strategi Penjualan Pada Toko Buku Gramedia Palembang Menggunakan Data Clustering*. Palembang: Jakarta.

Agustin, F. E., Fitria, A., & Hanifah, A. (2009). *Implementasi Algoritma K-Means Untuk Menentukan Kelompok Pengayaan Materi Mata Pelajaran Ujian Nasional (Studi Kasus: SMP Negeri 101 Jakarta)*. Jakarta: Universitas Islam Negeri Syarif Hidayatullah .