

Pengelompokan Data Menggunakan Algoritma K-Means Untuk Meningkatkan Motivasi Belajar Siswa MTS Alwashliyah 48 Kebun Lada

Anjuwita Situmorang¹, Yusfrizal², Indah Ambarita³

¹ STMIK KAPUTAMA ² UNIVERSITAS POTENSI UTAMA ³ STMIK KAPUTAMA
(e-mail: anjuwitasitumorang77@gmail.com, yusfrizal80@gmail.com, yesnovada@yahoo.com)

Abstrak

Belajar sudah menjadi kewajiban yang harus dilakukan oleh seorang siswa. Kadang siswa jenuh dan akhirnya memilih untuk tidak belajar untuk itulah mereka perlu diberi dorongan baik dari segala pihak. Pengaruh motivasi terhadap prestasi belajar siswa sangat besar karna itu hendaknya siswa diberi motivasi yang besar. Kedisiplinan dalam pengelolaan pengajaran merupakan suatu hal yang sangat penting. Tanpa adanya kesadaran akan keharusan melaksanakan aturan yang sudah ditentukan sebelumnya proses pengajaran tidak mungkin mencapai target yang maksimal. Penulisan proposal ini menggunakan metode *clustering* yang merupakan salah satu teknik data *Mining* untuk pengelompokan data siswa MTS Al-Wasliyah 48 Kebun Lada Binjai Dengan menggunakan metode *clustering* algoritma *k-means*. Dengan menerapkan 1101 sampel data alternatif data siswa dan memberikan jumlah *cluster* sebanyak 3, dan memanfaatkan 3 kriteria utama sebagai penelitian di proposal ini, menghasilkan jumlah *cluster* 1 sebanyak 578 data, *cluster* 2 sebanyak 256 data. Dan *cluster* 3 sebanyak 267 data. Sistem pengelompokan data siswa MTS Al-Wasliyah 48 binjai ini di rancang dengan Bahasa pemrograman aplikasi *MATLAB* dan memanfaatkan *Gui* sebagai *interfacenya*.

Kata kunci: Algoritma_K-Means, Clustering, Data_Mining, Data Siswa, MATLAB

Abstract

Learning has become an obligation that must be carried out by a student. Sometimes students get bored and in the end choose not to study because of that they need to be given good encouragement from all parties. The influence of motivation on learning achievement is very large because of that school students are given great motivation. Discipline in the management of teaching is a very important thing. Without awareness of the need to carry out predetermined rules, the teaching process will not be able to achieve the maximum target. Writing this proposal uses the clustering method which is a data mining technique for grouping student data at MTS Al-Wasliyah 48 Binjai Pepper Gardens using the k-means algorithm clustering method. By applying 1101 alternative data samples of student data and providing a total of 3 clusters, and utilizing 3 main criteria as research in this proposal, the resulting number of cluster 1 is 578 data, cluster 2 is 256 data. And cluster 3 as many as 267 data. The MTS Al-Wasliyah 48 binjai student data grouping system is designed with the MATLAB application programming language and utilizes the GUI as its interface.

Keywords: K-Means_Algorithm, Clustering, Data_Mining, Residual Data, MATLAB

1. Pendahuluan

Pendidikan sangat berperan penting sebagai proses peningkatan sebuah kualitas sumber daya manusia dalam upaya pembangunan sebuah bangsa dan Negara. Sekolah menjadi sarana berlangsungnya pendidikan secara langsung dengan bertemunya guru dan siswa. Guru merupakan faktor terpenting dalam proses peningkatan mutu pendidikan, selain itu keaktifan siswa juga merupakan unsur dasar yang penting bagi keberhasilan proses pembelajaran. Peningkatan mutu Pendidikan merupakan prioritas kebijakan umum di Indonesia yang harus terus diperhatikan. Belajar sudah menjadi kewajiban yang harus dilakukan oleh seorang siswa. Kadang siswa jenuh dan akhirnya memilih untuk tidak belajar untuk itulah mereka perlu diberi dorongan baik dari berbagai pihak. Pengaruhpenghasilan orang tua dan tingkat kehadiran terhadap prestasi belajar siswa sangat besar. Kedisiplinan dalam pengelolaan pengajaran merupakan suatu hal yang sangat penting. Tanpa adanya kesadaran akan keharusan melaksanakan aturan yang sudah ditentukan sebelumnya proses pengajaran tidak mungkin mencapai target yang maksimal. [1]

Kemampuan finansial yang berbeda-beda sedikit banyak akan berpengaruh terhadap prestasi belajar siswa. Status sosial ekonomi orang tua mewujudkan kemampuan finansial orang tua, tentunya akan mempengaruhi fasilitas belajar yang disediakan oleh orang tua terhadap sarana dan prasarana yang dibutuhkan oleh seorang siswa untuk meningkatkan prestasi belajarnya. Pada MTs Alwasliyah 48 Kebun Lada data siswa belum dikelola dengan baik dan data terus bertambah setiap tahunnya. Data tersebut masih kurang dimanfaatkan untuk menghasilkan suatu informasi.

Status sosial ekonomi orang tua siswa MTs Alwasliyah 48 Kebun Lada dan tingkat kedisiplinan bervariasi. Salah satu penyebab siswa malas bersekolah atau bahkan sampai tidak melanjutkan sekolah adalah faktor sosial dan ekonomi orang tua siswa. Data mining dibutuhkan untuk menghasilkan informasi yang berguna terhadap pihak MTs Alwashliyah 48 Kebun Lada khususnya untuk mengetahui informasi tentang prestasi belajar siswa berdasarkan status ekonomi dan prestasi belajar siswa..[2]

Dengan pemanfaatan data mining menggunakan data siswa, pada MTs Alwashliyah 48 Kebun Lada, data tersebut akan diolah dengan metode clustering menggunakan algoritma *K-Means* menggunakan variabel status sosial ekonomi siswa (penghasilan orang tua), variabel kedisiplinan (tingkat kehadiran siswa) dan variabel prestasi belajar (diambil dari nilai rata-rata raport siswa) pada MTs Alwashliyah 48 Kebun Lada. Berdasarkan hasil pengelompokan tersebut, pengolahan data yang dilakukan dapat menghasilkan informasi yang berguna terhadap pihak sekolah untuk mengetahui prestasi belajar siswa berdasarkan status ekonomi, tingkat kehadiran siswa dan nilai raportnya serta memudahkan pihak sekolah dalam upaya mengembangkan kemampuan yang dimiliki siswa.

Penelitian yang dilakukan oleh (Hasnan Afif et al.2022) dengan judul “Penerapan Data Mining Untuk Pengelompokan Kelas Siswa Unggulan Berdasarkan Nilai Raport Menggunakan Metode K-Means Clustering (Studi Kasus : SMK Muhammadiyah 3 Waleri)” menyatakan bahwa dalam pengelompokan kelas siswa unggulan menggunakan algoritma k-means dapat mengelompokkan siswa berdasarkan karakteristik yang relatif sama sehingga tidak terjadi kesenjangan sosial antara siswa yang memiliki kemampuan rendah dan siswa yang memiliki kemampuan tinggi sehingga proses pembelajaran dapat dilakukan dengan optimal dan dapat membantu pihak sekolah dalam meningkatkan prestasi siswa [3].

2. Metode Penelitian

Metode Penelitian memberikan penjelasan tentang langkah-langkah, data, lokasi penelitian, metode evaluasi yang digunakan serta penjelasan terstruktur tentang algoritma atau metode dari penelitian yang dibahas.



Gambar 1. Metodologi Penelitian

Metodologi penelitian ini dilakukan untuk mencari sesuatu secara sistematis dengan menggunakan metode ilmiah serta sumber yang berlaku. Dengan adanya proses ini, dapat memberikan hasil penelitian yang baik dan tepat. Atas dasar metodologi penelitian yang digunakan, maka dapat dibuat suatu alur kegiatan seperti dibawah ini :

1. Persiapan
Tahap ini adalah kegiatan awal dari penelitian, yaitu dengan penentuan dari latar belakang masalah kemudian dilakukan batasan masalah dan selanjutnya dilakukan penentuan tujuan serta manfaat yang dilakukan dalam penyusunan proses kerja sistem. Setelah itu, penulis menentukan kebutuhan data penelitian diantaranya menentukan metode yang akan digunakan, kriteria dari Data Siswa. Setelah itu penulis membuat surat izin penelitian yang ditujukan kepada Kepala Sekolah MTS Alwashliyah 48 Kebun Lada untuk meminta data dan mengumpulkan data yang dibutuhkan dalam penelitian ini.
2. Kajian Teori
Dalam tahap ini penulis mengumpulkan berbagai teori baik dari buku yang dipinjam dari perpustakaan, jurnal maupun internet untuk mendukung penelitian yang akan dilakukan. Teori yang dikumpulkan antara lain mengenai, Data Mining, Metode *Clustering*, algoritma *K-Means*, Data Siswa, Status Ekonomi, Penghasilan, Kedisiplinan, Tingkat Kehadiran, Prestasi Belajar, *Matlab*, *Gui*, *Flowchart* dan teori yang mendukung lainnya.
3. Pengumpulan Data
Tahap ini merupakan pengumpulan data yang diperlukan dalam pembuatan skripsi, yaitu Data Siswa yang diperoleh Dari MTs Alwashliyah 48 Kebun Lada
4. Analisa Data
Tahapan ini berupa merupakan tahapan mengelola dan menganalisa data yang telah diperoleh yaitu Data Siswa kemudian data tersebut dilakukan transformasi untuk dapat

dilakukan analisa yang kemudian diproses dengan menggunakan metode Clustering untuk mendapatkan sebuah informasi yang baru atau dikelompokkan sesuai dengan variabel yang telah ditentukan.

5. Perancangan Sistem

Tahap ini merupakan tahapan perancangan system yang akan dibangun dengan menggali data untuk mengelompokkannya dengan kriteria-kriteria (variabel) apa yang bisa dikelompokkan untuk dijadikan sebagai bahan penelitian. Adapun kriteria (variabel) yang diambil dari penelitian ini adalah Penghasilan Orang Tua, Kehadiran dan Nilai Raport siswa.

6. Pengujian dan implementasi

Tahap ini merupakan tahapan yang melakukan pengujian validasi dan implementasi data yang telah dianalisa sebelumnya serta penyusunan program. Pada tahap ini juga dilakukan pengujian dengan memproses data yang telah diproses dengan menggunakan *Software Matlab* sebagai implementasi proses pengelompokan Data Siswa berdasarkan status sosial ekonominya.

2.1. Data Mining dan Clustering

Data Mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan dalam *database*. *Data Mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan pembelajaran mesin (*machine learning*) untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan berpengetahuan yang terkait dari berbagai *database* besar [4].

Menurut (Cahyo Prianto et al. 2020) *clustering* adalah “suatu metode mempartisi data-set menjadi beberapa sub-set atau kelompok sedemikian rupa sehingga elemen-elemen dari suatu kelompok tertentu memiliki *set property* yang *dishare* bersama, dengan tingkat similaritas yang tinggi dalam suatu kelompok dan tingkat similaritas antar kelompok yang rendah”. Sehingga *cluster* tidak harus sama akan tetapi pengelompokannya berdasarkan pada kedekatan dari suatu karakteristik sampel yang ada salah satunya dengan menggunakan rumus jarak *eclidean*. Tujuan utama dari metode clustering adalah pengelompokan sejumlah data/obyek kedalam *cluster (group)* sehingga dalam setiap *cluster* akan berisi data yang semirip mungkin [5]

Menurut (Suhartini et al.2021) menyatakan bahwa *Data Mining* dapat menghasilkan informasi yang bermanfaat, dimana dapat dibuktikan bahwa hasil dari sistem yang dibuat dengan metode *clustering* yang dapat mengelompokkan data penduduk di wilayah Sukamulia Timur yang memang dikatakan tergolong penduduk miskin untuk memudahkan dalam proses pembagian bantuan agar merata kepada penduduk di wilayah Sukamulia Timur. Dengan menggunakan Algoritma *K-Means* dalam mengelompokkan data dapat dilakukan dengan sangat efektif. *Software* yang digunakan yaitu *rapid miner* dan *ms. Excel*. [7]

Menurut (Nurhayati.2022) menyatakan bahwa “ *K-Means* merupakan salah satu metode pengelompokan data non hierarki (sekatan) yang berusaha mempartisi data ke dalam *cluster* atau kelompok sehingga data yang memiliki karakteristik yang sama akan dikelompokkan ke dalam satu *cluster* yang sama dan data yang memiliki karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lain. [8]

2.2. MATLAB GUI

Menurut (Indah Werdiningsih et al.2022) menyatakan bahwa *Matlab* merupakan singkatan dari Matrix Laboratory. *Matlab* merupakan sebuah platfor pemogramana dan komputasi numeric yang digunakan untuk menganalisis data, mengembangkan algoritma dan membuat model. Leleh karena itu *matlab* banyak digunakan beberapa engineer dan ilmuwan di berbagai bidang riset yang memerlukan komputasi numeric yang komplek. [9]

Menurut (Tatik Retno Murniasih et al.2021) *GUIde* atau *GUI (Graphical User Interface)* adalah salah satu komponen dari *matlab* untuk membuat *interface (desain form)*

proses penyelesaian persoalan matematika yang lebih efisien dan menarik. Tidak seperti *m-file* hanya bisa bermain di *command windows*. [10]

3. Hasil dan Pembahasan

Langkah-langkah yang dilakukan untuk perhitungan data siswa menggunakan metode *clustering* dengan algoritma *k-means* agar dapat dihasilkan sebuah informasi baru tentang data siswa berdasarkan status sosial siswa, kedisiplinan siswa dan prestasi belajar siswa sehingga dapat diketahui hubungan terdekat antara kelompok data siswa.

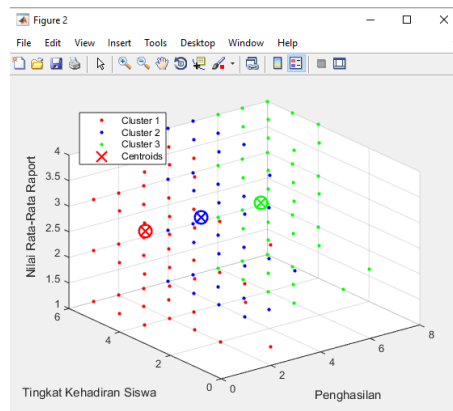
a. Contoh Perhitungan Clustering Menggunakan Algoritma K-Means

Adapun proses metode *clustering K-Means* berdasarkan data dari Sekolah MTS Alwashliyah 48 Kebun Lada Binjai adalah sebagai berikut :

Tabel 1. Perhitungan Hasil *Centroid*

Pendidikan (X)	Pekerjaan (Y)	Penghasilan (Z)	Keterangan
3,91	8,98	4,14	<i>Centroid 1</i>
27,06	20,41	9,13	<i>Centroid 2</i>
16,77	19,32	8,92	<i>Centroid 3</i>

Untuk lebih jelasnya hasil *cluster 3* dapat dilihat pada grafik dibawah ini :



Gambar 2. Grafik Hasil Perhitungan *Cluster 3* Berdasarkan Data Siswa

Pusatnya = 2 (3,91) 5 (8,98) 2(4,14)

 4(27,06) 5 (20,41) 2(9,13)

 7 (16,77) 5 (19,32) 2(8,92)

Keterangan :

Dari 1101 data siswa diperoleh 3 *cluster*, dimana *cluster 1* berjumlah 578 data , dan *cluster 2* berjumlah 256 data dan *cluster 3* berjumlah 267 data. Berikut adalah keterangan dari pusat *cluster* pada grafik :

1. 2 (3,91) 5 (8,98) 2(4,14)

Dapat diketahui bahwasanya pada *cluster 1* Kelompok data siswa dengan penghasilan orang tua Rp.1.000.001-Rp.1.500.000 dengan tingkat kehadiran siswa 81%-90% dan nilai rata raport 71-80

2. 4 (27,06) 5 (20,41) 2(9,13)

Dapat diketahui bahwasanya pada *cluster 2* Kelompok data siswa dengan penghasilan orang tua Rp.2.000.001-Rp.2.500.000 dengan tingkat kehadiran siswa 81%-90% dan nilai rata-rata raport 71-80

3. 7(16,77) 5 (19,32) 2(8,92)

Dapat diketahui bahwasanya pada *cluster 3* Kelompok data siswa dengan penghasilan orang tua Rp.3.500.001-Rp.4.000.000 dengan tingkat kehadiran siswa 81%-90% dan nilai rata raport 71-80



Gambar 3. Tampilan Form Menu Utama Pada Matlab

**PENGELOMPOKAN DATA SISWA BERDASARKAN STATUS SOSIAL EKONOMI, KEDISPLINAN DAN PRESTASI BELAJAR SISWA
MENGUNAKAN METODE CLUSTERING PADA MTS ALWASHLIYAH 48 KEBUN LADA**

Ambil Data

Jumlah Kolom:

Jumlah Baris:

Jumlah Cluster:

Proses

Menu Utama

Hasil Centroid

Centroid 1 :	2.26	5.19	2.40
Centroid 2 :	4.48	5.16	2.38
Centroid 3 :	6.93	5.23	2.34

Keterangan Hasil Centroid

2 - Rp 1.000.001-Rp1.500.000	5 = 81%-90%	2 = 71-80
4 - Rp 2.000.001-Rp 2.500.000	5 = 81%-90%	2 = 71-80
7 - Rp 3.500.001- Rp 4.000.000	5 = 81%-90%	2 = 71-80

Jarak Centroid

Jarak centroid 1 :	578	Data
Jarak centroid 2 :	267	Data
Jarak centroid 3 :	256	Data

Gambar 4. Tampilan Menu Proses Clustering Pada Matlab

4. Kesimpulan

Berdasarkan uraian bab-bab sebelumnya, maka dapat diambil beberapa kesimpulan penelitian ini antara lain sebagai berikut :

Dari hasil perhitungan algoritma K- Means dapat diketahui prestasi belajar siswa berdasarkan status ekonomi, kedisiplinan dan prestasi belajarnya. Dari hasil perhitungan algoritma K- Means dapat disimpulkan kelompok data siswa berprestasi paling banyak terdapat pada *cluster* 1 yaitu penghasilan orang tua siswa Rp. 1.000.001-Rp.1.500.000 dengan tingkat kehadiran siswa 81%-90% memiliki nilai rata-rata raportnya 71-80.

Dengan sistem yang dibangun dapat mempermudah pihak MTS Alwashliyah 48 Kebun Lada dalam mengambil keputusan dalam pemilihan siswa berprestasi dan dapat mempermudah pihak sekolah apabila ingin memberikan bantuan ataupun beasiswa. Serta membantu pihak sekolah dalam upaya mengembangkan kemampuan yang dimiliki siswa untuk meningkatkan prestasi belajarnya.

Berdasarkan kesimpulan di atas maka dikemukakan beberapa saran yang diharapkan dapat menjadi masukan bagi kemajuan sistem dimasa yang akan datang. Dan beberapa saran yaitu sebagai berikut :Perlu dilakukan penelitian lebih lanjut terhadap penggunaan algoritma *k-means* dan metode *clustering* yang lebih baik, yakni dengan akurasi yang lebih baik pula. Perlu dilakukan proses penggalan variasi variabel data agar hasil *clustering* yang dihasilkan dapat lebih maksimal.

Daftar Pustaka

- [1] M. Triandini, S. Defit, and G. W. Nurcahyo, "Data Mining dalam Mengukur Tingkat Keaktifan Siswa dalam Mengikuti Proses Belajar pada SMP IT Andalas Cendekia," *J. Inf. dan Teknol.*, vol. 3, pp. 167–173, 2021, doi: 10.37034/jidt.v3i3.120.
- [2] F. P. Dewi, P. S. Aryni, and Y. Umaidah, "Implementasi Algoritma K-Means Clustering Seleksi Siswa Berprestasi Berdasarkan Keaktifan dalam Proses Pembelajaran," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 7, no. 2, pp. 111–121, 2022, doi: 10.14421/jiska.2022.7.2.111-121.
- [3] R. P. Primanda, A. Alwi, and D. Mustikasari, "DATA MINING SELEKSI SISWA BERPRESTASI UNTUK MENENTUKAN KELAS UNGGULAN MENGGUNAKAN METODE K-MEANS CLUSTERING (Studi Kasus di MTS Darul Fikri)," *Komputek*, vol. 5, no. 1, p. 88, 2021, doi: 10.24269/jkt.v5i1.686.
- [4] W. S et al., *Data Mining*. Global Eksekutif Teknologi, 2023. [Online]. Available: <https://books.google.co.id/books?id=xmqvEAAAQBAJ>
- [5] C. Prianto and S. Bunyamin, *Pembuatan aplikasi clustering gangguan jaringan menggunakan metode K-Means clustering*. in Knowledge. Bandung.: Kreatif, 2020. [Online]. Available: <https://books.google.co.id/books?id=y8TgDwAAQBAJ>
- [6] S. Suhartini and R. Yuliani, "Penerapan Data Mining untuk Mengcluster Data Penduduk Miskin Menggunakan Algoritma K-Means di Dusun Bagik Endep Sukamulia Timur," *Infotek J. Inform. dan Teknol.*, vol. 4, no. 1, pp. 39–50, 2021, doi: 10.29408/jit.v4i1.2986.
- [7] S. Rizal and R. Q. Khotimah, "Penerapan Data Mining Untuk Clustering Data Penduduk Yang Terdampak Covid-19 Menggunakan Algoritma K-Means," *J. Pendidik. dan Konseling*, vol. 4, pp. 2781–2792, 2022.
- [8] NURHAYATI, *PEMODELAN K- MEANS ALGORITMA DAN BIG DATA ANALYSIS (PEMETAAN DATA MUSTAHIQ)*. Pascal Books, 2022. [Online]. Available: https://books.google.co.id/books?id=%5C_bJmEAAAQBAJ
- [9] I. Werdiningsih, D. C. R. Novitasari, and D. Z. Haq, *Pengelolaan Data Mining dengan*

Pemrograman Matlab. Airlangga University Press, 2022. [Online]. Available: <https://books.google.co.id/books?id=CgOdEAAQBAJ>

- [10] T. R. Murniasih, S. Suryanti, and V. Mandailina, *Tutorial Praktis Belajar GUI Matlab untuk Media Pembelajaran Matematika*. 2021.