

# Normalisasi Data Siswa Dengan Model Desimal Scaling Untuk Siswa Menerima Bantuan Dengan Algoritma K-Medoids

LENDRA CITRA DEFITRI

Fakultas Ilmu Komputer dan Teknologi Informasi, Program Studi Teknik Informatika, Universitas Budi Darma, Medan, Indonesia

Email: [lendracitra41@gmail.com](mailto:lendracitra41@gmail.com)

**Abstrak**— Program bantuan pendidikan sekolah merupakan salah satu program nasional dalam mengurangi biaya keperluan sekolah setiap siswa yang kurang mampu namun proses penyeleksian data pemilihan siswa yang layak di smp negeri 2 Panai Hulu masih tergolong penghitungan secara manual dengan data yang besar. Data mining merupakan metode perhitungan data dengan jumlah yang besar.

Menerapkan algoritma Data Mining pada data dengan rentang yang drastis dapat memberikan hasil yang kurang akurat. Karena itu dibutuhkan proses normalisasi data. Normalisasi data adalah salah satu teknik persiapan data yang paling sering digunakan pada Data Mining untuk unit, rentang, dan skala yang berbeda.

Normalisasi data hasil klasifikasi dapat lebih akurat karena semua kriteria data yang digunakan tidak saling mereduksi. Dalam mempermudah mengolah serangkaian informasi dalam Data Mining digunakan sebuah teknik pengelompokan yaitu Clustering pengelompokan berupa menyeleksi kriteria – kriteria seperti (pekerjaan orang tua, jumlah penghasilan orang tua, jumlah saudara kandung).

**Kata Kunci:** Bantuan; Normalisasi; Data Mining

**Abstract**— The school education assistance program is one of the national programs to reduce the cost of school supplies for each underprivileged student, but the process of selecting data for selecting eligible students at SMP Negeri 2 Panai Hulu is still considered manual calculations with large amounts of data. Data mining is a method of calculating large amounts of data.

Applying Data Mining algorithms to data with drastic ranges can produce less accurate results. Because of this, a data normalization process is needed. Data normalization is one of the most frequently used data preparation techniques in Data Mining for different units, ranges and scales.

Normalizing data from classification results can be more accurate because all the data criteria used do not reduce each other. To make it easier to process a series of information in Data Mining, a grouping technique is used, namely Clustering, in the form of selecting criteria such as (parents' occupation, amount of parents' income, number of siblings).

**Keywords:** Help; Normalization; Data Mining

## 1. PENDAHULUAN

Program Bantuan Pendidikan Sekolah merupakan salah satu program Nasional dalam mengurangi biaya keperluan sekolah, setiap siswa yang kurang mampu, sebagai bantuan yang halangani siswa miskin untuk mencegah putus sekolah dan siswa miskin yang kurang mampu mendapatkan kesempatan untuk duduk dibangku sekolah, tujuan dari program bantuan adalah agar siswa miskin mendapatkan pelayanan Pendidikan program pemerintah siswa wajib belajar 9 (sembilan) tahun. Bantuan Siswa Miskin (BSM) merupakan bentuk bantuan berupa uang yang diberikan pihak pemerintah secara langsung kepada siswa yang kurang mampu yang berasal dari keluarga miskin.

Sekolah Menengah Pertama Negeri Dua (SMPN 2) Panai Hulu merupakan sekolah yang menerapkan program pemerintah yaitu Bantuan Siswa Miskin (BSM). Dalam mengolah data siswa yang wajib memperoleh Bantuan Siswa Miskin (BSM), metode yang digunakan masih tergolong metode manual. Metode manual ini memiliki beberapa kekurangan, yaitu pada proses penyeleksian membutuhkan waktu yang lama, yaitu melakukan pendataan siswa yang akan memperoleh bantuan, menyeleksi kriteria seperti (pekerjaan orang tua siswa, penghasilan orang tua siswa, jumlah saudara kandung siswa, dan lainnya), dan mengelompokkan data siswa yang berhak dalam memperoleh Bantuan Siswa Miskin (BSM), sehingga harus melakukan proses mengelompokkan ulang untuk memperbaiki kesalahan-kesalahan dalam mengelompokkan data siswa.

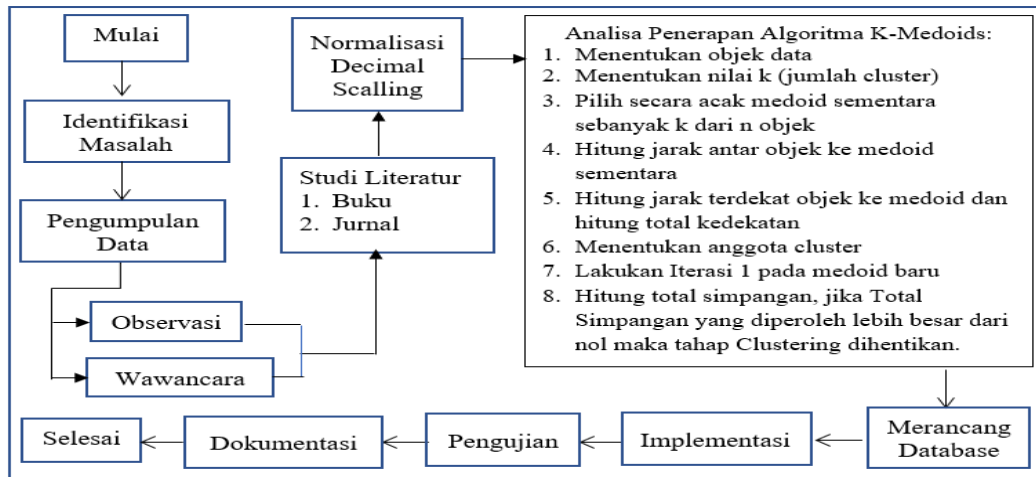
Masalah— masalah yang muncul selalu menjadi hambatan pada saat menerapkan program Bantuan Siswa Miskin (BSM) di SMPN 2 Panai Hulu. Pihak terkait SMPN 2 Panai Hulu. Dalam menciptakan efisiensi pengumpulan data yang besar tentunya diperlukan metode baru yang dapat menjawab kebutuhan tersebut. Data Mining merupakan sebuah teknologi yang dapat memproses data dalam volume besar yang digunakan oleh perusahaan untuk mengubah data mentah menjadi informasi yang berguna untuk membuat suatu keputusan bisnis yang sangat penting[1]. Data Mining merupakan penyederhanaan data dengan ukuran besar untuk mendapatkan serangkaian informasi yang bermanfaat dan mudah dipahami. Data Mining digunakan sebuah teknik pengelompokan data yang disebut dengan istilah Clustering. Tujuan Clustering adalah untuk mengelompokkan banyak data yang mempunyai beberapa kemiripan pada karakteristik untuk dimasukkan kedalam kelompok lainnya. Salah satu metode Clustering yang terdapat pada Data Mining yang digunakan untuk pengelompokan suatu data diantaranya adalah algoritma k-Medoids. Algoritma K-Medoids merupakan varian dari fungsi metode K-Means yang terdapat didalam Data Mining [2]. Algoritma K-Medoids adalah bagian dari metode partisi Clustering dengan menggabungkan atau mengelompokkan sekumpulan objek untuk mewakili Cluster.

Menerapkan algoritma Data Mining dengan data yang memiliki rentang drastis dapat menghasilkan data yang kurang tepat. Dalam proses seleksi data memerlukan normalisasi data. Model Decimal Scaling bekerja dengan membulatkan suatu bilangan desimal ke titik decimal yang paling terdekat, decimal scaling menormalisasikan dengan menggeser titik desimal dari angka.

## 2. METODOLOGI PENELITIAN

### 2.1 Kerangka Penelitian

Berikut ini adalah kerangka penelitian atau langkah pengumpulan informasi sampai pengujian dengan aplikasi RapidMiner:



Gambar 1 Kerangka Penelitian

Tahap penelitian dilakukan studi literatur dan studi lapangan. Studi literatur dilakukan untuk mengkaji dan mengetahui tahap kegiatan penelitian ini, dilakukan pemeriksaan terhadap objek yang akan diteliti dengan membaca beberapa sumber referensi seperti buku-buku, jurnal dan lain-lain. Sedangkan studi lapangan adalah mempelajari bagaimana metode yang sedang berjalan terkait objek penelitian di SMPN 2 Panai Hulu terletak di jalan Ajamu, Kecamatan Panai Hulu, Kabupaten Labuhanbatu. Tahap penelitian ini dilakukan berdasarkan pengumpulan data yang berupa observasi dan wawancara, studi dokumen.

### 2.2 Data Preprocessing

Data preprocessing merupakan suatu teknik yang pada umumnya diterapkan untuk database dengan menghapus beberapa data yang tidak konsisten, missing value, dan noise [3]. Adapun data ini terbagi menjadi tiga langkah, yaitu data cleaning, reduction, dan data transformation, dalam menggunakan data preprocessing ini data realtime database memiliki kelemahan seringkali tidak lengkap dan juga hasil data yang tidak konsisten sehingga mengakibatkan hasil akhir dari data mining kurang tepat dan tidak akurat. Untuk meningkatkan hasil pengolahan data yang tepat perlunya meningkatkan kualitas data yang akan dianalisis dengan melakukan langkah-langkah preprocessing data. Adapun langkah-langkah yang harus dilakukan peneliti ataupun data scientist yaitu perlunya mengenali langkah awal apa yang perlu dilakukan pada database yang peneliti miliki.

Data Cleaning, Hasil dari data yang telah dikumpulkan ada kemungkinan besar memiliki beberapa bagian yang hilang dan tidak relevan, sehingga perlu data cleaning untuk melaksanakan pembersihan data.

1. Transformation Data, Data transformation merupakan data yang digunakan untuk mengubah beberapa data ke bentuk yang tepat dalam proses data mining.
2. Normalisasi Data, Normalisasi data adalah elemen dasar data mining untuk memastikan record pada dataset tetap konsisten.
3. Teknik Normalisasi Data

### 2.3 Data Mining

Penambangan data (bahasa Inggris: data mining) adalah proses pencarian pengetahuan dari data dalam jumlah besar, menggunakan pengenalan pola, statistik dan teknik matematika. Data mining merupakan proses untuk menemukan pola data dan pengetahuan yang menarik dari kumpulan data yang sangat besar. Sumber data dapat mencakup database, data warehouse, web, repository, atau data yang dialirkan ke dalam sistem dinamis[4]. Penambangan data memiliki beberapa nama alternatif, meskipun definisi eksaknya berbeda, seperti KDD (knowledge discovery in database), analisis pola, arkeologi data, pemanenan informasi, dan intelegensia bisnis. data mining mempunyai hubungan dengan berbagai bidang diantaranya statistik, machine learning (pembelajaran mesin), pattern recognition, computing algorithms, database technology, dan highperformance computing [5].

Data mining mempunyai fungsi yang penting untuk membantu mendapatkan informasi yang berguna serta meningkatkan pengetahuan bagi pengguna[6]. Pada dasarnya, data mining mempunyai empat fungsi dasar yaitu: Fungsi Prediksi (prediction), Fungsi Deskripsi (description), Fungsi Klasifikasi (classification) dan Fungsi Asosiasi (association). Proses yang umumnya dilakukan oleh data mining antara lain: deskripsi, prediksi, estimasi, klasifikasi, clustering dan asosiasi. Tahapan yang dilakukan pada proses data mining diawali dari seleksi data dari data sumber ke data target, tahap preprocessing untuk memperbaiki kualitas data, transformasi, data mining serta tahap interpretasi dan evaluasi yang menghasilkan output berupa pengetahuan baru yang diharapkan memberikan kontribusi yang lebih baik.

## 2.4 K-Medoids

K-Medoid adalah teknik partisi klasik Clustering yang mengelompokkan data set dari ni objek ke dalam kelompok k yang dikenal apriori. Algoritma K-Medoids Clustering adalah salah satu algoritma yang digunakan untuk klasifikasi atau pengelompokan data. Algoritma K-Medoids ini, data akan digunakan sebagai nilai tengah/disebut dengan medoid, dan perhitungan jarak dihitung dari jarak antar masing-masing data [7]. Algoritma K-Medoids memiliki kelebihan untuk mengatasi kelemahan pada pada algoritma K-Means yang sensitif terhadap noise dan outlier, dimana objek dengan nilai yang besar yang memungkinkan menyimpang pada dari distribusi data. Kelebihan lainnya yaitu hasil proses Clustering tidak bergantung pada urutan masuk dataset [8]. K-Medoid merupakan kelompok metode partitionial Clustering yang meminimalkan jarak antara titik berlabel berada dalam cluster dan titik yang ditunjuk sebagai pusat kluster.

## 2.5 Program Bantuan Siswa Miskin

Program Bantuan Siswa Miskin (BSM) adalah bantuan yang diberikan kepada siswa dari keluarga kurang mampu untuk dapat melakukan kegiatan belajar di sekolah. Bantuan siswa miskin (BSM) merupakan salah satu strategi implementasi kebijakan pemerataan dan perluasan akses pendidikan yang bermutu bagi semua warga negara tanpa kecuali[9].

# 3. HASIL DAN PEMBAHASAN

Metode dalam menganalisis masalah penelitian ini yaitu menggunakan metode K-medoids, adapun Langkah – Langkah dalam metode K-medoids adalah sebagai berikut:

- Menentukan k (jumlah cluster) yang diinginkan, Ditentukan nilai  $K = 2$
- Pilih secara acak medoid awal sebanyak k dari n data. Berikut ini pemilihan objek secara acak yang terdiri dari objek 15 dan objek 16:

Objek	Pekerjaan Orang Tua	Jumlah Saudara Kandung	Penghasilan
15	0,000	0,000	237,500
16	0,000	0,000	100,000

Tabel 1 Pemilihan Objek Acak

- Hitung jarak masing-masing objek ke medoid sementara ke objek 1 dan objek dengan rumus Euclidean Distance:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

Hasil Pengujian K-Medoids diperoleh dari tahapan-tahapan Implementasi Aplikasi RapidMiner untuk Pengelompokan siswa yang berhak memperoleh bantuan Pendidikan (Studi kasus: SMPN 2(Dua) Panai Hulu. Adapun untuk melihat hasil Pengujian dari tahapan-tahapan Implementasi Aplikasi RapidMiner adalah dengan cara mengklik tombol Button Process, sehingga akan diperoleh hasil pengujian

## 3.1 Data Preprocessing

Data Preprocessing merupakan salah satu tahapan sebelum melakukan clustering dengan metode K-Medoids, tahapan – tahapan dalam Data Processing adalah sebagai berikut:

- Data Cleaning
- Transformasi Data
- Normalisasi data Model Decimal Scalling

## 3.2 Perancangan atau Pemodelan Sistem

Perancangan atau pemodelan sistem yang digunakan dalam analisa Normalisasi Data Siswa dengan Model Desimal Scaling untuk pengelompokan siswa yang berhak menerima bantuan pendidikan dengan Algoritma K-Medoids (Studi Kasus: SMPN 2 (Dua) Panai Hulu) terdiri dari, Activity Diagram, Activity Diagram Normalisasi data

siswa dengan Model Decimal Scalling untuk pengelompokan siswa yang layak menerima Bantuan Siswa Miskin (BSM) sebagai bantuan Pendidikan dengan Algoritma K-Medoids.

### 3.3 Implementasi

Tahap implementasi adalah tahap penerapan perancangan atau pemodelan sistem yang telah direncanakan sebelumnya yang bertujuan untuk mengelompokkan siswa yang berhak menerima bantuan pendidikan dengan Algoritma K-Medoids (Studi Kasus: SMPN 2 (Dua) Panai Hulu). Adapun tahap implementasi ini terdiri dari implementasi database dan implementasi Aplikasi.

#### a. Implementasi Database

Database yang akan di implementasikan bertipe dokumen titik dengan 20 sampel objek (data siswa). Berikut ini adalah Rancangan Database sekaligus Tabel yang akan dikelompokkan dengan Metode K-Medoids:

	A	B	C	D	E	F	G
	Pekerjaan Orang Tua	Jumlah Saudara Kandung	Penghasilan				
1							
2	0,000	0,001	150,000				
3	0,000	0,000	250,000				
4	0,000	0,000	100,000				
5	0,000	0,000	75,000				
6	0,000	0,001	87,500				
7	0,000	0,000	118,750				
8	0,000	0,000	250,000				
9	0,000	0,000	225,000				
10	0,000	0,000	187,500				
11	0,001	0,000	162,500				
12	0,000	0,001	112,500				
13	0,000	0,001	187,500				
14	0,000	0,000	225,000				
15	0,000	0,000	100,000				
16	0,000	0,000	237,500				
17	0,000	0,000	100,000				
18	0,000	0,000	81,250				
19	0,000	0,000	87,500				
20	0,000	0,001	175,000				
21	0,001	0,000	187,500				
22							
23							
24							

Gambar 2 Rancangan Database

#### b. Implementasi Aplikasi

Tahap Implementasi Aplikasi adalah tahap untuk proses pengujian Metode K-medoids berdasarkan Database yang telah dinormalisasi dengan Model Decimal Scalling. Tahap Implementasi Aplikasi ini menggunakan Aplikasi RapidMiner. Adapun Tahap Implementasi Aplikasi dengan Aplikasi RapidMiner adalah sebagai berikut:

1. Aktifkan Aplikasi RapidMiner
2. Dokumen baru
3. Import Data
4. Design

## 4. KESIMPULAN

Berdasarkan dari proses pengumpulan data, analisa, perancangan serta hasil pengujian dengan Aplikasi RapidMiner diperoleh beberapa kesimpulan sebagai berikut:

1. SMPN 2 (Dua) Panai Hulu merupakan sekolah yang menerapkan program pemerintah yaitu Bantuan Siswa Miskin (BSM) untuk membantu siswa yang berasal dari keluarga kurang mampu agar tidak putus sekolah.
2. Dalam mengelompokkan siswa yang berhak memperoleh Bantuan Siswa Miskin (BSM) pihak sekolah SMPN 2 (Dua) Panai Hulu masih menggunakan metode manual, karena metode manual yang memiliki banyak kelemahan diantaranya membutuhkan waktu yang relatif lama dalam proses pengelompokkan data siswa yang berhak mendapat Bantuan Siswa Miskin (BSM).
3. K-Medoids merupakan salah satu algoritma pengelompokkan data yang digunakan penulis untuk mengelompokkan data siswa yang berhak mendapat Bantuan Siswa Miskin (BSM). Algoritma K-Medoids ini menggunakan objek pada kumpulan objek untuk mewakili sebuah Cluster. Tujuan Clustering adalah untuk mengelompokkan data yang memiliki kesamaan karakteristik kedalam kelompok data yang sama dan data yang berbeda karakteristik kedalam kelompok yang lain.
4. Prosedur uji coba metode K-Medoids melalui Tool Aplikasi RapidMiner Studio 10.0 dalam mengelompokkan data siswa yang berhak menerima Bantuan Siswa Miskin (BSM) sebagai Bantuan Pendidikan di SMPN 2 (Dua) Panai Hulu.

5. Berdasarkan dengan perhitungan K-Medoids manual, proses pengujian K-Medoids untuk 20 items objek dengan aplikasi Rapid Miner juga memperoleh dua hasil cluster yang sama yaitu: Cluster 0 terdiri dari 9 items dan Cluster 1 terdiri dari 11 items.
6. Sesuai dengan perhitungan K-Medoids manual dan hasil Pengujian K-Medoids dengan aplikasi RapidMiner juga menunjukkan hasil yang sama yaitu Pusat Medoit terletak pada Objek ke 18 dan objek 19. Untuk objek 18 memiliki nilai kriteria pekerjaan orang tua (0,00), jumlah saudara kandung (0,00) dan penghasilan orang tua (87,00). Dan untuk objek 19 memiliki nilai kriteria pekerjaan orang tua (0,00), jumlah saudara kandung (0,01) dan penghasilan orang tua (187,500).

## UCAPAN TERIMAKASIH

Terima kasih disampaikan kepada pihak-pihak yang telah mendukung terlaksananya proses pengumpulan data dan sampai mendapatkan hasil akhir penelitian ini

## REFERENCES

- [1] Mustika & Dkk (2021). "Data Mining dan Aplikasinya", Penerbit Widina Bhakti Persada, Bandung, Edisi 3.
- [2] Zaenal Mustofa, Iman Saufik Suasana (2013). "Algoritma Clusteing K-Medoids Pada E-Government Bidang Information And CommunicationTechnology Dalam Menentukan Status Edge", Sistem Komputer STEKOM Semarang, Semarang.
- [3] JDQLab (2021). "Langkah-langkah Pengolahan Data", Yayasan Multimedia Nusantara & Xeratic, Retrieved from <https://dqlab.id/langkah-awal-dalam-pemrosesan-data-dalam-data-mining>.
- [4] Trivusi (2021). "Normalisasi Data: Pengertian, Tujuan, dan Metodenya", Retrieved from <https://www.trivusi.web.id/2022/09/normalisasi-data.html>.
- [5] M.R. Leksono & A.E. Prastiyano, penerapan algoritma Clustering untuk mengelompokan ketertaikan siswa berdasarkan aktifitas dimodel pembelajaran elektronik, Vol.1, 2014
- [6] S.Asmiratun, N. Wakhidah & A. N. Putri, penerapan K-Medoids untuk pengelompokan kondisi Jakarta di kota semarang, Vo.;6, No.2, Pp.176
- [7] F.Adam, Dkk, Perbandingan Normalisasi data untuk klasifikasi wine menggunakan algoritma naïve bayes, dicision tree dan support vector machine, Pp.264
- [8] Much Aziz & Dkk (2019). "Data Mining Algoritma C4.5", Universitas Negeri Semarang, Semarang.
- [9] F. L.Sibuea & A. Sapta (2017). Pemetaan Siswa Berprestasi Menggunakan Metode K-Means Clustering, Vol. 4, No. 1, Pp. 89
- [10] Gede Aditra & Ketut Agustini (2013). "Konsep Dasar Data Mining", Retrieved from <https://www.studocu.com/id/document/universitas-bandung-raya/information-technology/msim4403-m1-other/59550244>.
- [11] KajianPustaka.com (2017). "Pengertian, Fungsi, Proses dan Tahapan Data Mining", Retrieved from <https://www.kajianpustaka.com/2017/09/data-mining.html>.
- [12] Y. Mardi, Data mining:klasifikasi menggunakan algoritma c4.5, Vol.2, Pp.213 219
- [13] J. Eska, Penerapan data mining untuk prediksi penjualan wallpaper menggunakan algoritma c4.5, Vol. 2,No.2, Pp. 11, Mar.2016
- [14] Meiryani. (2019, Oct 03). Memahami Apa Itu Data Mining. Available: <https://accounting.binus.ac.id/201/10/03/memahami-apa-itu-data-mining>
- [15] Aditya Rahmatullah Pratama (2019). "Belajar Unified Modeling Language (UML) Pengenalan", CodePolitan.com, Retrieved from <https://codepolitan.com/blog/unified-modeling-language-uml>.
- [16] D. Septiani, P. D. Larasati, A. Irawan, Analisis Dan Perancangan Sistem Pengisian Kartu Rencana Study (KRS) Untuk Jurusan Teknik Informatika Dan Sistem Informasi Kampus Tanri Abeng University, Vol 1, No. 1. Pp. 22. 2018
- [17] Muhammad Fikry (2019). "Basis Data", Cetakan Pertama, Unimel Press, Lhokseumawe.
- [18] Deepublish (2023). "Pengertian dan Contoh Metodologi Penelitian", Retrieved from <https://deepublishstore.com/apa-itu-metodologi-penelitian/>
- [19] Hardani & Dkk (2020). "Buku Metode Penelitian Kualitatif & Kuantitatif ", Cetakan Pertama, CV. Pustakan Ilm Grup, Yogyakarta, Retrieved from <https://www.researchgate.net/publication/340021548>