

**ANALISIS *ALGORITMA K-MEANS* DALAM *CLUSTERING* SISWA
BERPRESTASI PADA PELAJARAN MATEMATIKA
(STUDI KASUS : SMA SWASTA KEMALA BHAYANGKARI
RANTAUPRAPAT)**

SKRIPSI

Untuk Memenuhi Persyaratan Memperoleh Gelar Sarjana (S1)
Program Studi Sistem Informasi Fakultas Sains Dan Teknologi
Universitas Labuhanbatu



OLEH :

AGATHA CHRISTIORENFA BR HALOHO

2109100139

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS LABUHANBATU
RANTAUPRAPAT**

2025

HALAMAN PERSEMBAHAN

Segala Puji dan Syukur kepada Tuhan Yang Maha Esa atas berkatnya kepada penulis, sehingga penulis dapat menyelesaikan skripsi ini tepat pada waktunya. Skripsi ini penulis persembahkan kepada:

1. Teristimewa kedua orang tua penulis, terima kasih atas segala doa dan dukungan yang tidak pernah putus. Memberikan cinta, kasih sayang, doa, dan pengorbanan yang mengiringi setiap langkah untuk menyelesaikan pendidikan ini. Terima kasih telah mengantarkan penulis sampai di titik ini. Terima kasih sudah berjuang, membesarkan dan mendidik penulis sampai mendapat gelar sarjana. Kiranya Tuhan senantiasa memberikan kesehatan, umur panjang dan mengiringi kehidupan kalian sampai melihat penulis berhasil dengan keputusan penulis sendiri.
2. Kepada saudaraku Fera Agrestina Lubis, Anaghella Christie Haloho, Amarin Chrishuld Haloho, dan Audrey Chrisluther Haloho, terima kasih telah menjadi bagian dari perjalanan hidup penulis. Telah menjadi energi positif dan tawa bahagia yang selalu kalian hadirkan. Semoga langkah kalian selalu diberikan perlindungan.
3. Teman seperjuangan penulis, Amanda Sartika Br Siregar dan Aida Amelia yang selalu membersamai serta membantu dalam kerumitan dalam menyusun skripsi penulis. Terimakasih sudah menjadi teman yang baik yang selalu memberikan motivasi, arahan dan semangat disaat penulis tidak percaya akan dirinya sendiri dan sempat hilang arah sehingga saat ini

penulis dapat menyelesaikan skripsi ini secara tepat waktu supaya dapat wisuda bersama-sama. Semoga Tuhan membalas segala kebaikan kalian.

4. Kepada seseorang yang tak kalah penting kehadirannya, pemilik NPM 1901100125. Terimakasih telah menjadi bagian dari perjalanan hidup penulis. Berkontribusi banyak dalam penulisan karya tulis ini, baik tenaga, waktu, maupun materi kepada penulis. Telah menjadi rumah, pendamping dalam segala hal yang menemani, mendukung ataupun menghibur dalam kesedihan, mendengar keluh kesah, memberi semangat untuk pantang menyerah. Semoga Tuhan selalu memberi berkat dalam segala hal yang kita lalui.
5. Serta semua pihak yang tidak dapat disebutkan satu per satu, yang telah banyak memberikan dukungan, motivasi, inspirasi dan membantu dalam proses penyelesaian skripsi ini.
6. Terakhir, terima kasih untuk diri sendiri karena sudah bertahan sejauh ini. Terima kasih tetap memilih berusaha dan merayakan dirimu sendiri sampai ada di titik ini, walau sering kali merasa putus asa atas apa yang diusahakan dan belum berhasil, namun terima kasih tetap menjadi manusia yang mau berusaha dan tidak lelah mencoba. Berbahagialah selalu dimanapun berada.

LEMBAR PERSETUJUAN SKRIPSI

JUDUL SKRIPSI : ANALISIS *ALGORITMA K-MEANS* DALAM
CLUSTERING SISWA BERPRESTASI PADA
PELAJARAN MATEMATIKA (STUDI KASUS :
SMA SWASTA KEMALA BHAYANGKARI
RANTAUPRAPAT)

NAMA : AGATHA CHRISTIORENEFA BR HALOHO

NPM : 2109100139

PRODI : SISTEM INFORMASI

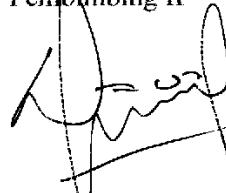
Disetujui pada tanggal : 01 September 2025

Pembimbing I



(Sudi Suryadi, S.Kom., M.Kom)
NIDN. 0128027903

Pembimbing II



(Gomal Juni Yahriz, S.Kom., M.Kom)
NIDN. 0124068301

LEMBAR PENGESAHAN SKRIPSI

JUDUL SKRIPSI : ANALISIS ALGORITMA K-MEANS DALAM
CLUSTERING SISWA BERPRESTASI PADA
PELAJARAN MATEMATIKA (STUDI KASUS :
SMA SWASTA KEMALA BHAYANGKARI
RANTAUPRAPAT)

NAMA : AGATHA CHRISTIORENFA BR HALOHO

NPM : 2109100139

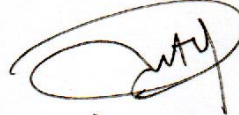
PRODI : SISTEM INFORMASI

Telah Diuji Dan Dinyatakan Lulus Dalam Uji Sarjana Pada
Tanggal 01..September 2025

TIM PENGUJI

Tanda Tangan

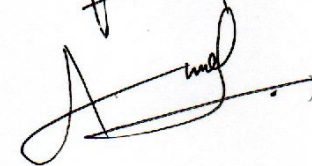
Pembimbing I (Ketua)
Nama : Sudi Suryadi, S.Kom., M.Kom
NIDN. 0128027903

()

Pembimbing II (Anggota)
Nama : Gomal Juni Yanris, S.Kom.,M.Kom
NIDN. 0124068301

()

Penguji (Anggota)
Nama : Angga Putra Juledi, S.Kom.,M.Kom
NIDN. 0119079401

()

Rantauprapat, 01..September 2025

Dekan
Fakultas Sains dan Teknologi

(Dediwan Purnama, S.Kom., M.Kom)
NIDN. 0112029202

Ka. Program Studi
Sistem Informasi

(Budianto Bangun, S.Sos., M.Kom)
NIDN. 0124047003

PERNYATAAN

JUDUL SKRIPSI : ANALISIS ALGORITMA K-MEANS DALAM
CLUSTERING SISWA BERPRESTASI PADA
PELAJARAN MATEMATIKA (STUDI KASUS :
SMA SWASTA KEMALA BHAYANGKARI
RANTAUPRAPAT)

NAMA : AGATHA CHRISTIORENFA BR HALOHO

NPM : 2109100139

PRODI : SISTEM INFORMASI

Dengan ini penulis menyatakan bahwa Skripsi ini disusun sebagai syarat untuk memperoleh gelar Sarjana pada Program Studi Sistem Informasi, Fakultas Sains dan Teknologi, Universitas Labuhanbatu adalah hasil karya tulis penulis sendiri. Semua kutipan maupun rujukan dalam penulisan Skripsi ini telah penulis cantumkan sumbernya dengan benar sesuai dengan ketentuan yang berlaku. Jika dikemudian hari ternyata ditemukan seluruh atau sebagian Skripsi ini bukan hasil karya penulis atau plagiat, penulis bersedia menerima sanksi pencabutan gelar akademik yang disandang dan sanksi-sanksi lainnya sesuai dengan peraturan Perundang-undangan yang berlaku.

Rantauprapat, 01 Sept. ... 2025

Yang Membuat Pernyataan



Agatha Christiorenfa Br Haloho

NPM. 2109100139

KATA PENGANTAR

Puji syukur saya panjatkan kepada Tuhan Yang Maha Esa, atas kehadirat-Nya, yang telah melimpahkan rahmat kepada saya, sehingga saya dapat menyelesaikan skripsi dengan judul “Analisis *Algoritma K-Means* Dalam *Clustering* Siswa Berprestasi Pada Pelajaran Matematika (Studi Kasus : SMA Swasta Kemala Bhayangkari Rantauprapat)“. Pada kesempatan ini, penulis hendak menyampaikan terima kasih yang sedalam – dalamnya kepada semua pihak yang telah memberikan dukungan moril maupun materil sehingga skripsi ini dapat diselesaikan. Ucapan terimakasih ini penulis tujukan kepada :

1. Bapak Halomoan Nasution, S.H., M.H selaku Ketua Yayasan Universitas Labuhanbatu
2. Bapak Assoc. Prof. Ade Parlaungan Nasution, Ph.D selaku Rektor Universitas Labuhanbatu,
3. Bapak Dr. Iwan Purnama, S.Kom., M.Kom selaku Dekan Fakultas Sains Dan Teknologi Universitas Labuhanbatu.
4. Bapak Budianto Bangun, S.Sos., M.Kom selaku Ketua Program Studi Sistem Informasi.
5. Bapak Sudi Suryadi, S.Kom., M.Kom Selaku Dosen Pembimbing I yang telah banyak membantu dan membimbing selama masa penyusunan skripsi ini.
6. Bapak Gomal Juni Yanris, S.Kom., M.Kom Selaku Dosen Pembimbing II yang telah membimbing, memberikan saran dan masukan hingga

terselesaikan nya skripsi ini.

7. Bapak Angga Putra Juledi, S.Kom., M.Kom selaku dosen penguji saya.
8. Seluruh Dosen Program Studi Sistem Informasi yang telah memberi bekal ilmu pengetahuan kepada penulis.

Penulis menyadari bahwa skripsi penelitian ini masih banyak kekurangan. Oleh karena itu penulis mengharapkan kritik dan saran yang membangun dari para pembaca guna menyempurnakan segala kekurangan dalam skripsi ini. Akhir kata, penulis berharap semoga skripsi ini berguna bagi para pembaca dan pihak-pihak lain yang berkepentingan.

Rantauprapat, 01 September 2025
Penulis



Agatha Christiorena Br Haloho
NPM. 2109100139

ABSTRAK

Penelitian ini bertujuan untuk menganalisis penerapan *algoritma K-Means Clustering* dalam mengelompokkan siswa berdasarkan prestasi belajar Matematika di SMA Swasta Kemala Bhayangkari Rantauprapat. Masalah utama yang dihadapi sekolah adalah belum adanya metode objektif untuk mengklasifikasikan siswa sesuai tingkat kemampuan akademik, sehingga diperlukan pendekatan berbasis data. Penelitian menggunakan data nilai harian, UTS, UAS, dan nilai akhir siswa kelas X, XI, dan XII tahun ajaran 2024/2025. Tahapan penelitian meliputi pengumpulan data, *preprocessing* (pembersihan, normalisasi, dan penanganan *missing value*), penerapan *algoritma K-Means*, evaluasi hasil kluster dengan *Davies-Bouldin Index* (DBI) dan *Silhouette Coefficient*, serta implementasi menggunakan perangkat lunak *RapidMiner*. Hasil analisis menunjukkan pembentukan tiga kluster: (1) siswa berprestasi tinggi dengan rata-rata nilai 88–95, (2) siswa berprestasi sedang dengan rata-rata nilai 80–85, dan (3) siswa berprestasi rendah dengan rata-rata nilai 70–76. Nilai DBI sebesar 0,182 menandakan kualitas kluster yang baik dengan tingkat pemisahan yang jelas antar kelompok. Temuan ini membuktikan bahwa *algoritma K-Means* efektif dalam memberikan segmentasi akademik yang akurat dan dapat menjadi dasar penyusunan strategi pembelajaran diferensiatif, seperti pemberian program pengayaan bagi siswa berprestasi tinggi serta pendampingan khusus bagi siswa berprestasi rendah. Penelitian ini juga menunjukkan potensi penerapan *data mining* dalam pendidikan sebagai alat bantu pengambilan keputusan berbasis analisis data yang sistematis.

Kata Kunci: *K-Means Clustering*, *Data Mining*, *RapidMiner*, Prestasi Belajar Matematika, Evaluasi Kluster.

ABSTRACT

This study aims to analyze the application of the K-Means Clustering algorithm in classifying students based on their Mathematics achievement at SMA Swasta Kemala Bhayangkari Rantauprapat. The main issue faced by the school is the absence of an objective method to categorize students according to their academic performance, making a data-driven approach necessary. The research uses data from daily assignments, midterm exams, final exams, and final Mathematics scores of students in grades X, XI, and XII for the 2024/2025 academic year. The research stages include data collection, preprocessing (data cleaning, normalization, and handling of missing values), implementation of the K-Means algorithm, cluster evaluation using the Davies-Bouldin Index (DBI) and Silhouette Coefficient, and system implementation with RapidMiner software. The analysis results show three distinct clusters: (1) high-achieving students with an average score of 88–95, (2) moderate-achieving students with an average score of 80–85, and (3) low-achieving students with an average score of 70–76. The DBI value of 0.182 indicates good cluster quality and clear separation among groups. These findings demonstrate that the K-Means algorithm effectively provides accurate academic segmentation and can serve as a basis for differentiated learning strategies, such as enrichment programs for high-achieving students and special assistance for low-achieving students. This research also highlights the potential of applying data mining in education as a systematic data-driven decision-making tool.

Keywords: K-Means Clustering, Data Mining, RapidMiner, Mathematics Achievement, Cluster Evaluation.

DAFTAR ISI

HALAMAN PERSEMBAHAN	ii
LEMBAR PERSETUJUAN SKRIPSI.....	iv
LEMBAR PENGESAHAN SKRIPSI	v
PERNYATAAN.....	vi
KATA PENGANTAR.....	vii
ABSTRAK	ix
<i>ABSTRACT</i>	x
DAFTAR ISI.....	xi
DAFTAR TABEL	xiii
DAFTAR GAMBAR.....	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Perumusan Masalah	2
1.3 Ruang Lingkup Masalah	2
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Tinjauan Umum Objek Penelitian.....	5
BAB II LANDASAN TEORI	7
2.1 <i>Data Mining</i> dalam Pendidikan.....	7
2.2 <i>Clustering</i>	9
2.3 <i>Algoritma K-Means</i>	11
2.4 Evaluasi Model <i>Clustering</i>	14
2.5 <i>RapidMiner</i>	16
2.6 Prestasi Belajar Matematika	20
BAB III METODOLOGI PENELITIAN	23
3.1 Kerangka Penelitian	23

3.2	Pengumpulan Data	25
3.3	<i>Preprocessing Data</i>	26
3.4	<i>Analisa Algoritma K-Means Clustering</i>	27
3.4.1	Menentukan <i>Cluster</i> dan Titik Pusat <i>Cluster</i>	27
3.4.2	Menghitung Jarak data terhadap <i>Centroid</i>	28
3.4.3	Hasil Perhitungan Manual Metode <i>K-Means</i>	38
BAB IV IMPLEMENTASI DAN PEMBAHASAN		41
4.1	Hasil Implementasi <i>RapidMiner</i>	41
4.2	Pembahasan	48
BAB V KESIMPULAN DAN SARAN		49
5.1	Kesimpulan	49
5.2	Saran	50
DAFTAR PUSTAKA		52

DAFTAR TABEL

Tabel 2. 1 Perbandingan <i>Hierarchical</i> dan <i>Partitional Clustering</i>	10
Tabel 2. 2 Interpretasi Nilai DBI.....	15
Tabel 2. 3 Interpretasi Nilai <i>Silhouette Score</i>	16
Tabel 3. 1 Data Sampel Nilai Siswa Kelas X – XII	26
Tabel 3. 2 Data Set Nilai Siswa.....	26
Tabel 3. 3 Jumlah <i>Cluster</i> Yang Digunakan.....	27
Tabel 3. 4 <i>Centroid</i> Awal.....	28
Tabel 3. 5 Jarak data terhadap <i>Centroid</i> Awal.....	28
Tabel 3. 6 <i>Cluster</i> Pada Iterasi Pertama	29
Tabel 3. 7 <i>Centroid</i> Iterasi Kedua	30
Tabel 3. 8 Jarak Data Pada Iterasi Kedua.....	30
Tabel 3. 9 <i>Cluster</i> Pada Iterasi Kedua.....	31
Tabel 3. 10 Perbandingan <i>Cluster</i> Iterasi Pertama dan Iterasi Kedua.....	31
Tabel 3. 11 <i>Centroid</i> Iterasi Ketiga	32
Tabel 3. 12 Jarak Data Pada Iterasi Ketiga	32
Tabel 3. 13 <i>Cluster</i> Pada Iterasi Ketiga.....	32
Tabel 3. 14 Perbandingan <i>Cluster</i> Iterasi Kedua dan Iterasi Ketiga	33
Tabel 3. 15 <i>Centroid</i> Iterasi Keempat	34
Tabel 3. 16 Jarak Data Pada Iterasi Keempat.....	34
Tabel 3. 17 <i>Cluster</i> Pada Iterasi Keempat.....	34
Tabel 3. 18 Perbandingan <i>Cluster</i> Iterasi Ketiga dan Iterasi Keempat	35
Tabel 3. 19 <i>Centroid</i> Iterasi Kelima	35

Tabel 3. 20 Jarak Data Pada Iterasi Kelima	36
Tabel 3. 21 <i>Cluster</i> Pada Iterasi Kelima	36
Tabel 3. 22 Perbandingan <i>Cluster</i> Iterasi Keempat dan Iterasi Kelima	37
Tabel 3. 23 <i>Cluster</i> 1	38
Tabel 3. 24 <i>Cluster</i> 2	38
Tabel 3. 25 <i>Cluster</i> 3	39
Tabel 4. 1 Deskripsi data berdasarkan <i>cluster</i>	47

DAFTAR GAMBAR

Gambar 2. 1 Proses Umum <i>Data Mining</i>	9
Gambar 2. 2 Diagram Alir Metode <i>K-Means</i>	13
Gambar 3. 1 Kerangka Kerja Penelitian	23
Gambar 4. 1 Memilih Lembar Kerja Awal	40
Gambar 4. 2 Lembar Kerja <i>RapidMiner</i> dan Menu Import Data	41
Gambar 4. 3 Proses Import Data.....	41
Gambar 4. 4 Input Operator Multiply	42
Gambar 4. 5 Input Operator <i>Algoritma K-Means</i>	42
Gambar 4. 6 Alur Proses <i>RapidMiner</i>	43
Gambar 4. 7 Hasil <i>Clusterisasi</i> dan Hasil Nilai DBI	44
Gambar 4. 8 Hasil <i>Cluster 0</i>	44
Gambar 4. 9 Hasil <i>Cluster 1</i>	45
Gambar 4. 10 Hasil <i>Cluster 2</i>	45
Gambar 4. 11 Visualisasi <i>Clustering</i>	46
Gambar 4. 12 <i>Centroid Table</i>	46