

**PENERAPAN ALGORITMA K-MEANS CLUSTERING UNTUK
PEMBENTUKAN ROMBONGAN BELAJAR BERDASARKAN
HASIL BELAJAR SISWA**

SKRIPSI

Oleh:
RISKA DWI ANGGRAENI
NIM. 18650111



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2024**

**PENERAPAN ALGORITMA K-MEANS CLUSTERING UNTUK
PEMBENTUKAN ROMBONGAN BELAJAR BERDASARKAN
HASIL BELAJAR SISWA**

SKRIPSI

Diajukan kepada:
Universitas Islam Negeri Maulana Malik Ibrahim Malang
Untuk Memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Sarjana Komputer (S.Kom)

Oleh:
RISKA DWI ANGGRAENI
NIM. 18650111

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2024**

HALAMAN PERSETUJUAN

**PENERAPAN ALGORITMA K-MEANS CLUSTERING UNTUK
PEMBENTUKAN ROMBONGAN BELAJAR BERDASARKAN
HASIL BELAJAR SISWA**

SKRIPSI

**Oleh :
RISKA DWI ANGGRAENI
NIM. 18650111**

Telah Diperiksa dan Disetujui untuk Diuji:
Tanggal: 02 Desember 2024

Pembimbing I,



H. Fatchurrohman, M.Kom
NIP. 19700731 200501 1 002

Pembimbing II,



Prof. Dr. Suhartono, S.Si, M.Kom
NIP. 19680519 200312 1 001

Mengetahui,
Ketua Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang



Dr. Fachrul Kurniawan, M.MT, IPU
NIP. 19771020 200912 1 001

HALAMAN PENGESAHAN

PENERAPAN ALGORITMA K-MEANS CLUSTERING UNTUK PEMBENTUKAN ROMBONGAN BELAJAR BERDASARKAN HASIL BELAJAR SISWA

SKRIPSI

Oleh :
RISKA DWI ANGGRAENI
NIM. 1865011

Telah Dipertahankan di Depan Dewan Penguji Skripsi
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
Untuk Memperoleh Gelar Sarjana Komputer (S.Kom)
Tanggal: 13 Desember 2024

Susunan Dewan Penguji

Ketua Penguji : Dr. Zainal Abidin, M.Kom
NIP. 19760613 200501 1 004

Anggota Penguji I : Supriyono, M.Kom
NIP. 19841010 201903 1 012

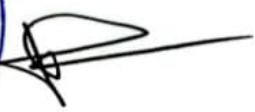
Anggota Penguji II : H. Fatchurrohman, M.Kom
NIP. 19700731 200501 1 002

Anggota Penguji III : Prof. Dr. Suhartono, S.Si, M.Kom
NIP. 19680519 200312 1 001

()
()
()
()

Mengetahui dan Mengesahkan,
Ketua Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang




Dwi Pachrul Kurniawan, M.MT, IPU
NIP. 19771020 200912 1 001

PERNYATAAN KEASLIAN TULISAN

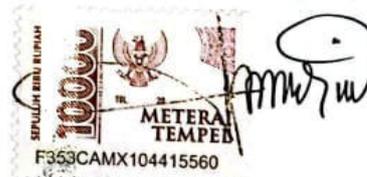
Saya yang bertanda tangan di bawah ini:

Nama : Riska Dwi Anggraeni
NIM : 18650111
Fakultas / Program Studi : Sains dan Teknologi / Teknik Informatika
Judul Skripsi : Penentu Rombongan Belajar Berdasarkan Hasil Belajar Siswa Menggunakan Algoritma K-Means Clustering

Menyatakan dengan sebenarnya bahwa Skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambil alihan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila dikemudian hari terbukti atau dapat dibuktikan skripsi ini merupakan hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 02 Desember 2024
Yang membuat pernyataan,



Riska Dwi Anggraeni
NIM.18650111

HALAMAN MOTTO

لَا يُكَلِّفُ اللَّهُ نَفْسًا إِلَّا وُسْعَهَا

“Allah tidak membebani seseorang, kecuali menurut kesanggupannya”

HALAMAN PERSEMBAHAN

Skripsi ini saya persembahkan untuk:

Kedua orang tua saya, Adik,

Keluarga, Sahabat,

Dan untuk diri saya sendiri

Terimakasih

KATA PENGANTAR

Assalamua'laikum Warahmatullahi Wabarakatuh

Puji dan syukur atas kehadiran Tuhan semesta alam Allah SWT yang telah memberikan Rahmat dan hidayah-Nya, sehingga peneliti diberikan kemudahan dan keberkahan dalam setiap menyelesaikan skripsi ini. Penyusunan skripsi ini bertujuan untuk memenuhi syarat kelulusan bagi mahasiswa Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang. Keberhasilan penulisan skripsi ini tidak lepas dari dorongan dan bimbingan dari berbagai pihak. Untuk itu dalam kesempatan ini penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Prof. Dr. H.M. Zainuddin, MA selaku rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Prof. Dr. Hj. Sri Harini, M.Si selaku dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Dr. Fachrul Kurniawan, M.MT, IPM selaku ketua Program Studi Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang.
4. H. Fatchurrahman, M.Kom dan Prof. Dr. Suhartono, S.Si, M.Kom selaku dosen pembimbing I dan II yang telah meluangkan waktu dalam membimbing dan memberikan arahan serta dorongan kepada peneliti.
5. Dr. Zainal Abidin, M.Kom selaku dosen penguji I dan Supriyono, M.Kom selaku dosen penguji II yang telah bersedia meluangkan waktu untuk memberikan pengarahan yang mendalam serta saran untuk pengembangan skripsi ini.

6. Kedua orang tua, Bapak Saipul Anwar dan Ibu Nuriamah yang mendukung, mendoakan, serta memberikan semangat kepada penulis dalam menyelesaikan skripsi ini. Adik tercinta Nadin Angelina Putri yang memberi semangat. Serta Dzakiyah Farhana Archila A.
7. Seluruh dosen dan staff program studi Teknik Informatika yang telah memberikan ilmu dan pengalaman yang berharga.
8. Sahabat dan teman-teman baik saya
9. Kepala madrasah, seluruh dewan guru MI Miftahul Uquiyah dan wali kelas 3 Ibu Ni'matul Mufidah yang selalu memberi semangat dalam menyelesaikan skripsi
10. Alisa Lila Nidha, Fauzan Adzhim, dan rekan kerja di Badan Adhoc yang kebersamai penulis menyelesaikan skripsi ini
11. Diri saya sendiri yang mampu untuk bertahan dan terus maju walau banyak kendala dan halangan di tengah-tengah pengerjaan skripsi ini

Penulis menyadari bahwa penyusunan skripsi ini jauh dari kata sempurna. Oleh sebab itu penulis mengharapkan kritik dan saran yang bersifat konstruktif. Penulis berharap bahwa terdapat manfaat yang bisa diambil dari skripsi ini.

Malang, 02 Desember 2024

Penulis

DAFTAR ISI

HALAMAN PENGAJUAN	ii
HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN	iv
PERNYATAAN KEASLIAN TULISAN	v
HALAMAN MOTTO	vi
HALAMAN PERSEMBAHAN	vii
KATA PENGANTAR	viii
DAFTAR ISI	x
DAFTAR GAMBAR	xii
ABSTRAK	xiv
ABSTRACT	xv
مستخلص البحث	xvi
BAB I PENDAHULUAN	1
1.1 Latar Belakang Penelitian	1
1.2 Pernyataan Masalah	4
1.3 Tujuan Penelitian	5
1.4 Manfaat Penelitian	5
1.5 Batasan Penelitian	6
1.6 Sistematika Penelitian	7
BAB II STUDI PUSTAKA	9
2.1 Konsep Dasar Data Mining	9
2.2 Konsep Dasar Algoritma <i>K-means</i>	11
2.2.1 Definisi Algoritma <i>K-means</i>	11
2.2.2 Proses Algoritma <i>K-means</i>	12
2.2.3 Kelebihan dan Kelemahan Algoritma <i>K-means</i>	13
2.3 Penerapan Algoritma <i>K-means</i> dalam Sistem Penilaian Hasil Belajar	16
2.3.1 Penggunaan Algoritma <i>K-means</i> dalam Penilaian Hasil Belajar.....	16
2.3.2 Pemilihan jumlah <i>cluster</i> yang tepat.....	17
2.3.3 Karakteristik data siswa yang kompleks	18
2.4 Nilai-nilai Islam dalam Pendidikan.....	18
2.4.1 Peran Pengetahuan dalam Islam.....	18
2.4.2 Penerapan Teknologi dalam Sistem Penilaian Hasil Belajar Siswa dalam Konteks Islam.....	19
2.5 Landasan Teori.....	19
2.5.1 Teori Pembelajaran	20
2.5.2 Teori Evaluasi Pendidikan	20
2.5.3 Teori Pemrosesan Data dan Analisis	20
2.6 Hubungan antara variabel penelitian.....	21
BAB III DESAIN DAN IMPLEMENTASI	23
3.1 Data	23
3.2 Proses K-Means Clustering	27
3.3 Implementasi	34

BAB IV HASIL DAN PEMBAHASAN.....	36
4.1 Hasil Penelitian	36
4.1.1 Uji Coba ke-1	38
4.1.2 Uji Coba ke-2	40
4.1.3 Uji Coba ke-3	42
4.1.4 Uji Coba ke-4	44
4.1.5 Uji Coba ke-5	46
4.1.6 Uji Coba Ke-6.....	48
4.2 Pembahasan.....	53
BAB V KESIMPULAN DAN SARAN	57
5.1 Kesimpulan	57
5.2 Saran.....	57
DAFTAR PUSTAKA	

DAFTAR GAMBAR

Gambar 2. 1 Bagian dari KDD.....	11
Gambar 2. 2 Flowchart Algoritma K-Means	12
Gambar 3. 1 grafik elbow method.....	27
Gambar 3. 2 Scatterplot.....	34

DAFTAR TABEL

Tabel 3. 1 Data nilai siswa	24
Tabel 3. 2 Parameter data.....	28
Tabel 3. 3 Hasil perhitungan manual Algoritma K-Means	30
Tabel 4. 1 Data nilai siswa	36
Tabel 4. 2 Data uji.....	38
Tabel 4. 3 Hasil uji coba 1	38
Tabel 4. 4 Hasil uji coba 2	41
Tabel 4. 5 Hasil uji coba 3	42
Tabel 4. 6 Hasil uji coba 4 iterasi ke-4.....	44
Tabel 4. 7 Hasil uji coba 5	47
Tabel 4. 8 Hasil uji coba 6	49
Tabel 4. 9 Hasil Silhouette coefficient.....	51
Tabel 4. 10 Hasil seluruh uji coba.....	51

ABSTRAK

Anggraeni, Riska Dwi. 2024. **PENENTU ROMBONGAN BELAJAR BERDASARKAN HASIL BELAJAR SISWA MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING**. Skripsi. Program Studi Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Fatchurrahman, M.Kom (II) Prof. Dr. Suhartono,S.Si, M.Kom

Kata Kunci : *K-Means Clustering, Elbow Method, Silhouette Score*

Penilaian hasil belajar siswa merupakan faktor untuk meningkatkan proses Pendidikan dan mengidentifikasi kebutuhan pengembangan lebih lanjut. Penelitian ini bertujuan untuk mengelompokkan siswa kedalam tiga kelas sesuai dengan hasil yang di hitung dengan algoritma k-means. Data pada penelitian ini terdiri dari 54 data siswa. Dari hasil perhitungan pada seluruh siswa mendapatkan hasil sebanyak 16 siswa dengan kelompok pemahaman tinggi, 19 siswa dengan pemahaman sedang dan 19 siswa dengan pemahaman kurang. Hasil dari silhouette score tertinggi adalah 0,5901 dan terendah adalah 0,5800.

ABSTRACT

Anggraeni, Riska Dwi. 2024. **PENENTU ROMBONGAN BELAJAR BERDASARKAN HASIL BELAJAR SISWA MENGGUNAKAN ALGORITMA *K-MEANS CLUSTERING***. Thesis. Informatics Engineering Study Program, Faculty of Science and Technology, Maulana Malik Ibrahim State Islamic University Malang. Supervisor: (I) Fatchurrahman, M.Kom (II) Prof. Dr. Suhartono,S.Si, M.Kom

Assessment of student learning outcomes is a factor to improve the education process and identify further development needs. This research aims to group students into three classes according to the results calculated by the k-means algorithm. The data in this study consisted of 54 student data. From the results of calculations on all students get results as many as 16 students with high understanding groups, 19 students with moderate understanding and 19 students with less understanding. The result of the highest silhouette score is 0.5901 and the lowest is 0.5800.

Keywords : *K-Means Clustering, Elbow Method, Silhouette Score*

مستخلص البحث

أنغريني، ريسكا دوي. 2024. تحديد مجموعات الدراسة بناء على مخرجات تعلم الطلاب باستخدام خوارزمية تجميع كي-متوسط. البحث الجامعي. قسم الهندسة المعلوماتية، كلية العلوم والتكنولوجيا بجامعة مولانا مالك إبراهيم الإسلامية الحكومية مالانج. المشرف الأول: فتح الرحمن، الماجستير. المشرف الثاني: أ. د. سوهارتونو، الماجستير.

الكلمات الرئيسية: تجميع كي-متوسط، طريقة لإيجاد عدد عناقيد، درجة صورة ظليلة

يعد تحسين نتائج تعلم الطلاب عاملا في تحسين العملية التعليمية وتحديد المزيد من احتياجات التنمية. هدف هذا البحث إلى تجميع الطلاب في ثلاثة فصول وفقا للنتائج المحسوبة بواسطة خوارزمية تجميع كي-متوسط. تألفت البيانات في هذا البحث من 54 بيانات للطلاب. من نتائج الحساب على جميع الطلاب، كان هناك 16 طالبا من مجموعات الفهم العالية، و 19 طالبا ذوي الفهم المتوسط و 19 طالبا ذوي الفهم المنخفض. نتيجة أعلى درجة الصورة الظلية هي 0.5901 وأقلها 0.5800.

BAB I

PENDAHULUAN

1.1 Latar Belakang Penelitian

Indonesia sebagai negara berkembang dengan populasi yang besar, menghadapi tantangan yang signifikan dalam meningkatkan kualitas dan akses terhadap pendidikan bagi seluruh warga negaranya. Pendidikan merupakan salah satu aspek penting dalam kehidupan manusia, yang memiliki peran besar dalam membentuk karakter dan potensi seseorang (Macintyre dkk., 2019). Dalam konteks ini, penilaian hasil belajar siswa menjadi faktor krusial untuk memastikan efektivitas proses pendidikan dan mengidentifikasi kebutuhan pengembangan lebih lanjut. Di era teknologi informasi yang semakin maju, pemanfaatan algoritma dan teknik-teknik analisis data dalam sistem penilaian hasil belajar siswa memiliki potensi yang besar untuk meningkatkan kualitas Pendidikan (Kuleto dkk., 2021).

Meskipun metode penilaian konvensional seperti ulangan harian, ujian tengah semester, dan ujian akhir semester masih banyak digunakan, penelitian menunjukkan bahwa metode-metode ini seringkali tidak efisien dan kurang akurat dalam mengidentifikasi kebutuhan individu siswa (Ferawati Jafar, 2021). Selain itu, persaingan global dan perkembangan teknologi informasi membutuhkan sistem pendidikan yang lebih adaptif dan responsif terhadap kebutuhan siswa (Szymkowiak dkk., 2021).

Penting untuk mengidentifikasi siswa sejak dini, dengan memberikan perhatian untuk mencapai retensi siswa (Syahrani & Zaman, 2023). MI Miftahul Uquiyah adalah salah satu institusi pendidikan Islam yang bertujuan untuk

meningkatkan kualitas pendidikan dan menciptakan lingkungan belajar yang baik bagi para siswanya. Kelas 3 di MI Miftahul Uquiyah menjadi fokus dalam penelitian ini karena pada tahap ini, siswa mulai menghadapi berbagai tantangan dalam proses pembelajaran yang memerlukan penilaian hasil belajar yang efektif dan efisien.

Salah satu pendekatan dalam analisis data adalah clustering, yang merupakan teknik pengelompokan data berdasarkan karakteristik dan pola-pola tertentu. Algoritma *K-means* merupakan salah satu algoritma clustering yang populer dan efisien dalam mengelompokkan data. Algoritma ini digunakan untuk mengelompokkan data yang tidak berlabel (*unsupervised*) ke dalam sejumlah kelompok (*cluster*) yang ditentukan berdasarkan kemiripan fitur atau atribut (Gangga dkk., 2014). *K-means* bekerja dengan menghitung jarak antara data dan menempatkan data yang mirip dalam satu kelompok. Penerapan algoritma *K-means* dalam menentukan rombongan belajar berdasarkan hasil belajar siswa di kelas 3 MI Miftahul Uquiyah dapat membantu guru dan institusi pendidikan dalam mengidentifikasi pola-pola prestasi belajar siswa, sehingga dapat meningkatkan pemahaman mengenai kebutuhan dan potensi siswa secara lebih efektif.

Dalam konteks Islam, Alqur'an menyampaikan konsep penting mengenai pendidikan dan pengetahuan, yang dapat memberikan pandangan dalam memahami pentingnya penerapan teknologi dalam sistem penilaian hasil belajar siswa. Alqur'an mengajarkan bahwa manusia harus menjelajahi dan mempelajari dunia untuk memperoleh pengetahuan (Al-'Alaq: 1-5). Selain itu Allah menjelaskan bahwa Allah menjanjikan peningkatan derajat bagi mereka yang beriman dan

memiliki ilmu. Ini menekankan pentingnya mencari ilmu dan bertindak berdasarkan iman serta dorongan untuk terus belajar. Seperti pada salah satu ayat Al-Qur'an Surat Al-Mujadallah Ayat 11:

يَا أَيُّهَا الَّذِينَ آمَنُوا إِذَا قِيلَ لَكُمْ تَفَسَّحُوا فِي الْمَجَالِسِ فَافْسَحُوا يَفْسَحِ اللَّهُ لَكُمْ ۗ وَإِذَا قِيلَ انشُرُوا فَانشُرُوا
يَرْفَعِ اللَّهُ الَّذِينَ آمَنُوا مِنْكُمْ وَالَّذِينَ أُوتُوا الْعِلْمَ دَرَجَاتٍ ۗ وَاللَّهُ بِمَا تَعْمَلُونَ خَبِيرٌ

“Wahai orang-orang yang beriman! Apabila dikatakan kepadamu, 'Berlapang-lapanglah dalam majelis', maka lapangkanlah, niscaya Allah akan memberi kelapangan untukmu. Dan apabila dikatakan, 'Berdirilah kamu', maka berdirilah, niscaya Allah akan meninggikan orang-orang yang beriman di antaramu dan orang-orang yang diberi ilmu beberapa derajat. Dan Allah Maha Teliti apa yang kamu kerjakan.”(Q.S Al Mujadillah: 11).

Dalam kitab tafsir *Al-Munir*, Allah memberikan penghargaan dengan mengangkat derajat kepada orang-orang yang beriman dan berilmu. Allah SWT secara khusus mengangkat derajat orang-orang yang memadukan iman dan ilmu. Mereka akan mendapatkan keistimewaan ini, karena Allah SWT meninggikan posisi mereka berkat keimanan dan pengetahuan yang dimiliki.

Ayat ini menunjukkan bahwa Allah SWT memberikan kedudukan tinggi kepada orang-orang yang beriman dan berilmu. Algoritma K-Means digunakan untuk mengelompokkan siswa berdasarkan kemampuan mereka. Dengan klusterisasi, sekolah dapat memberikan pendekatan pembelajaran yang lebih spesifik sesuai dengan kebutuhan tiap kluster. Siswa dengan kemampuan rendah diberikan perhatian khusus untuk meningkatkan pemahaman mereka, sehingga kualitas pendidikan meningkat. Penerapan algoritma *K-means* dalam sistem penilaian hasil belajar siswa di MI Miftahul Uquliyah tidak hanya sejalan dengan

kemajuan teknologi, tetapi juga mencerminkan nilai-nilai Islam dalam meningkatkan kualitas pendidikan.

Namun, penerapan clustering dan algoritma *K-means* untuk penentuan rombongan belajar hasil belajar siswa di MI Miftahul Uquiyah masih memiliki beberapa tantangan, seperti pemilihan jumlah *cluster* yang tepat, penentuan metode evaluasi yang efektif, dan pemahaman mengenai karakteristik data siswa yang kompleks. Oleh karena itu, perlu adanya penelitian yang mendalam untuk mengatasi masalah-masalah ini dan mengoptimalkan penerapan algoritma K-means dalam sistem penilaian hasil belajar siswa di MI Miftahul Uquiyah.

Berdasarkan pemikiran teoritis dan fakta empirik yang telah dijelaskan, terdapat kesenjangan atau "gap" dalam penerapan *clustering* dan algoritma *K-means* pada sistem penilaian hasil belajar siswa di MI Miftahul Uquiyah. Oleh karena itu, penelitian dengan judul "Penentu Rombongan Belajar Berdasarkan Hasil Belajar Siswa Menggunakan Algoritma K-Means *Clustering*" menjadi penting dan relevan untuk diteliti, dengan tujuan untuk meningkatkan efektivitas dan efisiensi sistem penilaian hasil belajar siswa di kelas 3 MI Miftahul Uquiyah.

1.2 Pernyataan Masalah

Berdasarkan latar belakang penelitian yang telah dijelaskan, beberapa pernyataan masalah yang muncul dalam penelitian ini adalah: Bagaimana mengimplementasikan algoritma *K-means* dalam penilaian hasil belajar siswa di kelas 3 MI Miftahul Uquiyah agar dapat mengidentifikasi pola-pola prestasi belajar siswa?

1.3 Tujuan Penelitian

Berdasarkan pernyataan masalah yang telah dipilih, tujuan penelitian ini adalah untuk mengimplementasikan algoritma *K-means* dalam penilaian hasil belajar siswa di kelas 3 MI Miftahul Uquiyah untuk mengelompokkan data prestasi belajar siswa berdasarkan karakteristik dan pola-pola tertentu.

1.4 Manfaat Penelitian

Penelitian ini diharapkan memberikan manfaat sebagai berikut:

1. Bagi institusi pendidikan (MI Miftahul Uquiyah) penelitian ini dapat membantu institusi dalam meningkatkan penilaian hasil belajar siswa dengan mengidentifikasi pola-pola prestasi belajar menggunakan algoritma *K-means*. Dengan demikian, institusi dapat mengambil keputusan yang tepat dalam proses pembelajaran dan mengalokasikan sumber daya dengan lebih efisien.
2. Bagi guru penelitian ini dapat membantu guru dalam memahami kebutuhan dan potensi siswa dengan lebih baik, sehingga dapat menyesuaikan metode pengajaran dan strategi pembelajaran sesuai dengan karakteristik siswa. Hal ini akan meningkatkan efektivitas proses pembelajaran dan membantu guru dalam mencapai tujuan pendidikan yang diharapkan.
3. Bagi siswa penelitian ini dapat memberikan informasi mengenai pola-pola prestasi belajar, yang dapat membantu siswa dalam meningkatkan pemahaman tentang kekuatan dan kelemahan mereka serta mengembangkan strategi belajar.
4. Bagi peneliti dan akademisi penelitian ini dapat menjadi referensi dan dasar penelitian lebih lanjut mengenai penerapan algoritma *K-means* dalam penilaian hasil belajar siswa, terutama di konteks institusi pendidikan berbasis Islam. Hal

ini akan memperkaya literatur ilmiah dalam bidang teknologi informasi dan pendidikan.

5. Bagi pengembang sistem penilaian dan teknologi pendidikan penelitian ini dapat memberikan wawasan dan inspirasi dalam mengembangkan sistem penilaian yang lebih efisien dengan menggunakan algoritma *K-means* dan teknologi informasi lainnya, sehingga dapat meningkatkan kualitas dan relevansi sistem penilaian hasil belajar siswa.

1.5 Batasan Penelitian

Dalam penelitian ini, terdapat beberapa batasan yang perlu diperhatikan untuk memfokuskan penelitian dan menghindari ambiguitas dalam proses penelitian:

- a. Fokus penelitian: Penelitian ini akan fokus pada penerapan algoritma *K-means* dalam sistem penilaian hasil belajar siswa, dengan tujuan pola-pola prestasi belajar siswa di kelas 3 MI Miftahul Uquiyah,
- b. Data penelitian: Penelitian ini akan menggunakan data hasil belajar siswa kelas 3 MI Miftahul Uquiyah, termasuk nilai ujian, tugas, dan kegiatan lainnya yang relevan. Data demografis siswa dan informasi lain yang tidak berkaitan langsung dengan prestasi belajar tidak akan dimasukkan dalam analisis,
- c. Metode analisis: Penelitian ini akan menggunakan algoritma *K-means* sebagai metode analisis utama. Meskipun ada berbagai metode *clustering* lain yang dapat digunakan, penelitian ini akan terbatas pada penerapan algoritma *K-means* karena efisiensinya dan popularitasnya dalam berbagai aplikasi,

- d. Ruang lingkup penelitian: Penelitian ini akan terbatas pada peningkatan penilaian hasil belajar siswa di kelas 3 MI Miftahul Uquliyah. Aspek-aspek lain yang berkaitan dengan sistem pendidikan, seperti pengajaran, kurikulum, dan manajemen sekolah, tidak akan dibahas dalam penelitian ini.

1.6 Sistematika Penelitian

Sistematika penelitian ini terdiri dari lima bab, yang dijelaskan sebagai berikut:

BAB I: PENDAHULUAN: Bab ini mencakup latar belakang penelitian yang menjelaskan pentingnya topik penelitian dan relevansinya, pernyataan masalah yang menggambarkan pertanyaan penelitian, tujuan penelitian yang menjelaskan apa yang diharapkan dicapai, hipotesis yang merupakan dugaan awal terkait pertanyaan penelitian, manfaat penelitian yang menjelaskan kontribusi penelitian ini untuk berbagai pihak, batasan penelitian yang menjelaskan fokus dan keterbatasan penelitian, serta sistematika penelitian yang menjelaskan struktur penulisan penelitian ini.

BAB II: TINJAUAN PUSTAKA: Bab ini berisi landasan teori yang menjelaskan teori-teori dasar yang relevan dengan penelitian, penelitian terdahulu yang membahas topik serupa, dan kerangka berpikir yang menggambarkan hubungan antara variabel penelitian. Tinjauan pustaka ini bertujuan untuk memberikan konteks dan pemahaman yang lebih dalam tentang topik penelitian.

BAB III: DESAIN DAN IMPLEMENTASI: Bab ini menjelaskan metode penelitian yang digunakan dalam penelitian ini, termasuk pola dan rancangan penelitian yang menjelaskan desain penelitian secara keseluruhan, bahan atau

materi penelitian yang mencakup data yang digunakan, alat yang digunakan dalam penelitian seperti perangkat keras dan perangkat lunak, serta langkah-langkah penelitian yang menjelaskan proses penelitian secara rinci.

BAB IV: HASIL DAN ANALISIS DATA: Bab ini akan memaparkan hasil penelitian, termasuk pengumpulan data, pra-pemrosesan data, implementasi algoritma *K-means*, analisis pola prestasi belajar siswa, evaluasi efektivitas penerapan algoritma *K-means*, dan pembahasan hasil penelitian.

BAB V: KESIMPULAN DAN SARAN: Bab ini berisi kesimpulan yang merangkum temuan penelitian dan menjawab pertanyaan penelitian, serta saran untuk penelitian lebih lanjut dan penerapan hasil penelitian dalam praktik.

BAB II

STUDI PUSTAKA

2.1 Konsep Dasar Data Mining

Secara garis besar, data mining dapat dikelompokkan menjadi dua kategori utama (Sofiani & Retno Utari, 2022), yaitu:

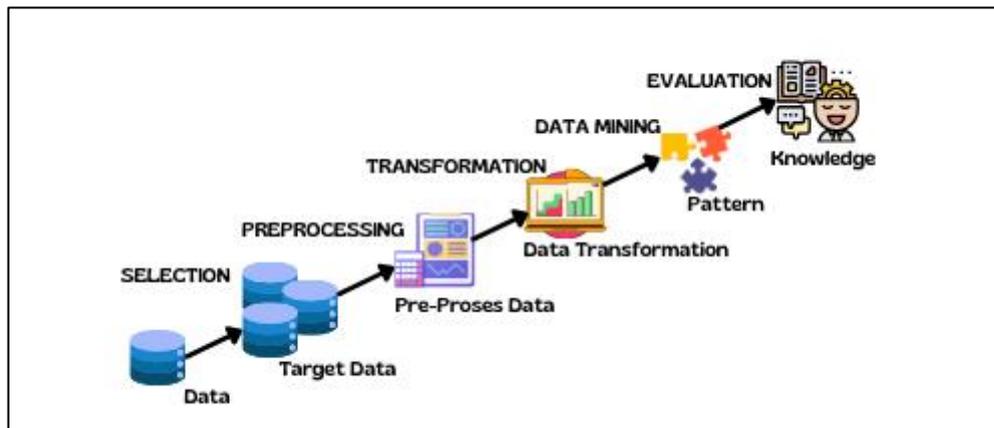
Pertama adalah data mining deskriptif: Kategori ini berfokus pada pengidentifikasian pola dan hubungan dalam data yang menjelaskan karakteristik dasar dan perilaku data. Tujuan utama data mining deskriptif adalah untuk menyajikan informasi yang berguna dan mudah dipahami tentang data. Beberapa teknik yang umum digunakan dalam data mining deskriptif meliputi klustering, analisis asosiasi, dan analisis urutan.

- 1) Klustering: Merupakan teknik yang digunakan untuk mengelompokkan objek atau titik data yang serupa ke dalam kelompok atau *cluster*. Algoritma klustering populer meliputi *K-means*, *hierarchical clustering*, dan DBSCAN.
- 2) Analisis asosiasi: Teknik ini digunakan untuk mengidentifikasi hubungan antara item atau variabel dalam data. Salah satu contoh yang terkenal adalah aturan asosiasi dalam analisis keranjang belanja (*market basket analysis*) untuk mengidentifikasi kombinasi produk yang sering dibeli bersama-sama.
- 3) Analisis urutan: Merupakan teknik yang digunakan untuk mengidentifikasi pola atau urutan yang sering muncul dalam data yang diurutkan, seperti urutan klik dalam log situs web atau urutan pembelian produk dalam transaksi pelanggan.

Kedua adalah data mining prediktif: Kategori ini berfokus pada pembuatan model yang dapat digunakan untuk memprediksi hasil atau perilaku di masa depan berdasarkan data historis. Data mining prediktif sering digunakan dalam pengambilan keputusan, analisis risiko, dan peramalan. Beberapa teknik yang umum digunakan dalam data mining prediktif meliputi klasifikasi, regresi, dan analisis rangkaian waktu.

- a. Klasifikasi: Merupakan teknik yang digunakan untuk memprediksi kategori atau label untuk objek atau titik data berdasarkan atribut mereka. Algoritma klasifikasi populer meliputi *decision trees*, *support vector machines (SVM)*, dan *neural networks*.
- b. Regresi: Teknik ini digunakan untuk memprediksi nilai kontinu atau numerik berdasarkan atribut objek atau titik data. Beberapa metode regresi yang umum digunakan meliputi regresi linier, regresi logistik, dan regresi polinomial.
- c. Analisis rangkaian waktu: Merupakan teknik yang digunakan untuk memprediksi nilai di masa depan berdasarkan data waktu yang telah terjadi sebelumnya. Metode analisis rangkaian waktu yang populer meliputi *ARIMA*, *Exponential Smoothing*, dan *LSTM (Long Short-Term Memory)* dalam *deep learning*.

Istilah lain dari Data Mining adalah *Knowledge Discovery in Database (KDD)* (Komariyah dkk., 2023). Walaupun data mining sendiri adalah bagian dari KDD. Seperti yang terlihat pada gambar 1 seperti berikut.



Gambar 2. 1 Bagian dari KDD

2.2 Konsep Dasar Algoritma *K-means*

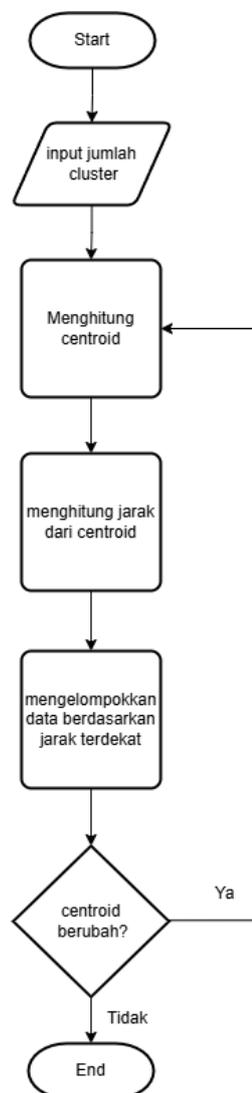
Algoritma *K-Means* merupakan salah satu metode dalam analisis *clustering* yang paling sering digunakan untuk mengelompokkan data ke dalam beberapa kluster berdasarkan kemiripan atribut. *K-Means* bekerja dengan cara membagi data menjadi sejumlah k kluster, di mana k adalah parameter yang telah ditentukan sebelumnya oleh pengguna.

2.2.1 Definisi Algoritma *K-means*

Algoritma *K-means* adalah sebuah teknik pengelompokan data atau klustering yang merupakan bagian dari metode *unsupervised learning* dalam bidang *machine learning* dan *data mining* (Ganga dkk., 2014). Algoritma ini bertujuan untuk mengelompokkan objek atau data yang memiliki karakteristik serupa ke dalam k *cluster* yang telah ditentukan sebelumnya (Jain, Murty, & Flynn, 1999). Salah satu ciri khas dari algoritma *K-means* adalah kemudahannya dalam implementasi serta efisiensi dalam mengelompokkan data dalam jumlah besar.

2.2.2 Proses Algoritma *K-means*

Secara umum langkah yang digunakan dalam *clustering* dengan metode *k-means* dimulai dari mengumpulkan data yang akan dikelompokkan. Jika data sudah terkumpul jumlah klaster dan centroid awal dapat ditentukan. Kemudian jarak dari centroid dihitung sampai dengan centroid tidak berubah. Flowchart dari algoritma *k-means* ditunjukkan oleh gambar berikut:



Gambar 2. 2 Flowchart Algoritma K-Means

Proses algoritma *K-means* melibatkan beberapa langkah utama, seperti dijelaskan oleh MacQueen (Pérez-Ortega dkk., t.t.), yaitu:

- 1) Inisialisasi: Menentukan jumlah *k-cluster* yang diinginkan dan menginisialisasi *centroid* acak untuk setiap *cluster*.
- 2) Penugasan: Menghitung jarak antara setiap titik data dan *centroid*, lalu menugaskan setiap titik data ke *cluster* yang memiliki *centroid* terdekat.

Proses perhitungan jarak menggunakan persamaan *Euclidean distance* :

$$d(x_i v_c) = x_i - v_c = \sqrt{\sum_{j=1}^m (x_{ij} - v_{cj})^2} \quad (2.1)$$

keterangan :

$d(x_i v_c)$: jarak dari data ke-*i* ke *centroid* ke-*c*
 x_{ij} : data ke-*i* dan atribut ke-*j*
 v_{cj} : *centroid* ke-*c* dan sribut ke-*j*

- 3) Pembaruan: Menghitung ulang *centroid* untuk setiap *cluster* dengan mengambil rata-rata dari titik-titik yang termasuk dalam *cluster* tersebut.

$$v_{cj} = \frac{\sum_{x_i \in \text{Clust}_c} x_{ij}}{|\text{Clust}_c|} \quad (2.2)$$

dimana:

$x_i \in \text{Clust}_c$: data ke-*i* pada *cluster* ke-*c*
 $|\text{Clust}_c|$: jumlah data pada *cluster* ke-*c*
 x_{ij} : data ke-*i* dan atribut ke-*j*
 v_{cj} : *centroid* ke-*c* dan atribut ke-*j*

- 4) Pengulangan: Mengulangi langkah 2 dan 3 hingga *centroid* tidak berubah atau jumlah iterasi yang telah ditentukan tercapai.

2.2.3 Kelebihan dan Kelemahan Algoritma *K-means*

Kelebihan algoritma *K-means* meliputi (Yuan & Yang, 2019):

- a. Mudah untuk diimplementasikan dan diinterpretasikan.

- b. Efisien dalam hal waktu komputasi, terutama untuk data dalam jumlah besar.
- c. Bisa digunakan untuk mengelompokkan data dalam berbagai bidang, seperti analisis citra, pengenalan pola, dan sistem rekomendasi.

Namun, algoritma *K-means* juga memiliki beberapa kelemahan (Yuan & Yang, 2019):

- a. Sensitif terhadap inisialisasi centroid awal, yang bisa menyebabkan hasil yang berbeda pada setiap eksekusi.
- b. Jumlah *k cluster* harus ditentukan sebelumnya, yang bisa menjadi tantangan jika tidak ada pengetahuan sebelumnya tentang struktur data.
- c. Tidak efektif untuk mengelompokkan data yang memiliki bentuk atau densitas yang tidak seragam.

Dalam penelitian ini, algoritma *K-means* akan diaplikasikan pada penilaian hasil belajar siswa di kelas 3 MI Miftahul Uquiyah, dengan mempertimbangkan kelebihan dan kelemahan algoritma ini. Sumber-sumber yang digunakan dalam pembahasan ini berasal dari referensi primer, seperti jurnal penelitian dan buku teks, yang memiliki relevansi serta keterkaitan dengan masalah yang dibahas dalam penelitian.

Algoritma *K-means* dan Data Mining sering digunakan secara bersamaan dalam proyek data mining untuk menggali informasi yang berguna dan membangun model yang dapat membantu pengambilan keputusan yang lebih efektif dan efisien (Lia Hananto dkk., 2021).

Data mining adalah proses ekstraksi pengetahuan atau informasi yang berguna dari kumpulan data besar melalui penggunaan berbagai teknik dan

algoritma. Salah satu algoritma yang sering digunakan dalam data mining adalah algoritma *K-means*. Algoritma *K-means* merupakan teknik klustering yang masuk ke dalam kategori data mining deskriptif (Khairani & Sutoyo, 2020). Data Mining adalah alat yang kuat untuk memberikan informasi dari data yang besar dan kompleks.

Hubungan antara data mining dan algoritma *K-means* terletak pada tujuan penggunaan *K-means* dalam proses data mining. Dalam data mining, *K-means* digunakan untuk mengidentifikasi struktur atau pola yang tersembunyi dalam data dengan mengelompokkan objek atau titik data yang serupa ke dalam kelompok (klaster) berdasarkan fitur-fitur yang dimilikinya. Algoritma *K-means* membagi data menjadi *K klaster* yang telah ditentukan sebelumnya, dengan setiap klaster direpresentasikan oleh *centroid* atau pusat klaster.

Algoritma *K-means* bekerja dengan cara menginisialisasi *K* titik pusat klaster secara acak, kemudian mengelompokkan titik data ke dalam klaster terdekat berdasarkan jarak (misalnya jarak *Euclidean*) antara titik data dan titik pusat klaster. Setelah semua titik data telah dikelompokkan, titik pusat klaster diperbarui dengan menghitung rata-rata dari titik-titik data yang termasuk dalam klaster tersebut. Proses ini diulangi hingga titik pusat klaster tidak berubah lagi atau jumlah iterasi yang ditentukan telah dicapai.

Dalam konteks data mining, algoritma *K-means* dapat digunakan untuk berbagai tujuan, seperti segmentasi pasar, analisis pola perilaku, deteksi anomali, dan pengurangan dimensi (Maria dkk., 2022). Dengan mengelompokkan data menjadi klaster yang homogen, *K-means* membantu mengungkap pola dan struktur

yang mendasari data, yang kemudian dapat digunakan untuk menginformasikan keputusan bisnis atau penelitian lebih lanjut.

2.3 Penerapan Algoritma *K-means* dalam Sistem Penilaian Hasil Belajar

Penerapan algoritma K-Means dalam sistem penilaian hasil belajar siswa bertujuan untuk mengelompokkan siswa berdasarkan kemampuan mereka dalam bidang studi tertentu. Dengan mengimplementasikan K-Means, sekolah dapat lebih mudah menganalisis performa siswa dan mengidentifikasi kelompok-kelompok dengan karakteristik hasil belajar yang serupa, sehingga dapat diterapkan strategi pembelajaran yang lebih tepat sasaran.

2.3.1 Penggunaan Algoritma *K-means* dalam Penilaian Hasil Belajar

Sistem Evaluasi pembelajaran saat ini merupakan sistem evaluasi konvensional. Sistem ini merupakan sistem tradisional yang umum dalam evaluasi pembelajaran. Sistem evaluasi ini menggunakan ujian tertulis untuk mengetahui pemahaman seorang siswa terhadap materi yang sudah dipelajari. Selain itu dalam sistem ini parameter yang digunakan untuk penilaian kepada seluruh siswa sama sehingga hal ini dianggap kurang objektif dalam proses evaluasi. Berbeda dengan sistem evaluasi menggunakan metode K-Means yang dilakukan dengan mengelompokkan siswa berdasarkan kesamaan karakteristik yang kemudian proses penilaian didasarkan pada parameter yang sesuai karakteristik kelompok tersebut dengan menggunakan kriteria-kriteria untuk penilaiannya.

Algoritma *K-means* dapat digunakan dalam sistem penilaian hasil belajar siswa untuk mengelompokkan siswa berdasarkan pola-pola prestasi belajar yang

serupa. Dalam konteks ini, data siswa seperti nilai ujian, kehadiran, dan partisipasi dalam kegiatan ekstrakurikuler dapat digunakan sebagai input untuk algoritma *K-means*. Pengelompokan siswa ini dapat membantu guru dan institusi pendidikan dalam mengidentifikasi kebutuhan dan potensi siswa, serta mengambil keputusan yang tepat dalam proses pembelajaran (Ganga dkk., 2014; Pérez-Ortega dkk., t.t.).

2.3.2 Pemilihan jumlah *cluster* yang tepat

Salah satu tantangan dalam penerapan algoritma *K-means* adalah menentukan jumlah *cluster* yang tepat. Beberapa metode yang dapat digunakan untuk memilih jumlah *cluster* yang optimal meliputi metode *Elbow*, *Silhouette*, dan *Gap Statistic* (Sagala & Gunawan, 2022). Setiap metode memiliki kelebihan dan kelemahan masing-masing, sehingga penting untuk mempertimbangkan karakteristik data dan tujuan penelitian saat memilih metode yang paling sesuai.

Elbow Method adalah teknik yang digunakan untuk menentukan jumlah kluster optimal dalam algoritma K-Means. Pendekatan ini berdasarkan analisis nilai inerti, yaitu jumlah kuadrat jarak antara setiap titik data dengan centroid kluster terdekat. Cara menentukan *elbow method* adalah dengan mencari titik di grafik dengan penurunan inerti yang mulai melambat. Jumlah kluster pada titik ini dianggap optimal.

Silhouette Score adalah metrik yang digunakan untuk mengevaluasi kualitas klusterisasi dengan mengukur seberapa baik titik data dikelompokkan. *Silhouette score* digunakan untuk mengevaluasi dan memilih jumlah kluster optimal, terutama jika metode *elbow* tidak memberikan hasil yang jelas.

2.3.3 Karakteristik data siswa yang kompleks

Data siswa yang digunakan dalam penelitian ini mungkin memiliki karakteristik yang kompleks, seperti adanya pencilan (*outliers*), atribut yang berkorelasi, dan variabilitas yang tinggi. Dalam konteks ini, penting untuk melakukan pra-pemrosesan data, seperti normalisasi, transformasi, dan seleksi fitur, sebelum menerapkan algoritma *K-means* (Ganga dkk., 2014). Selain itu, mungkin perlu untuk mempertimbangkan penggunaan algoritma klustering lain yang lebih cocok untuk mengelompokkan data siswa yang kompleks.

2.4 Nilai-nilai Islam dalam Pendidikan

Nilai-nilai Islam dalam pendidikan adalah prinsip-prinsip dan pedoman yang diambil dari ajaran agama Islam yang dijadikan dasar dalam proses pendidikan. Nilai-nilai ini tidak hanya berkaitan dengan aspek akademis tetapi juga membentuk karakter, moral, dan perilaku peserta didik agar menjadi pribadi yang berakhlak mulia dan beriman kepada Allah SWT. Penerapan nilai-nilai Islam dalam pendidikan bertujuan untuk menciptakan manusia yang seimbang antara pengetahuan duniawi dan spiritual, serta bertanggung jawab sebagai khalifah di muka bumi.

2.4.1 Peran Pengetahuan dalam Islam

Dalam pandangan Islam, pengetahuan merupakan bagian penting dari kehidupan seseorang, yang diperlukan untuk mengembangkan pemahaman tentang tujuan hidup dan menjalani kehidupan yang lebih baik. Dalam Hadis, Nabi Muhammad SAW mengatakan, "Menuntut ilmu adalah kewajiban bagi setiap

Muslim." (Ibn Majah). Pengetahuan dalam Islam tidak hanya mencakup agama, tetapi juga meliputi ilmu-ilmu duniawi, seperti sains, teknologi, dan seni.

2.4.2 Penerapan Teknologi dalam Sistem Penilaian Hasil Belajar Siswa dalam Konteks Islam

Penerapan teknologi dalam sistem penilaian hasil belajar siswa, seperti penggunaan algoritma *K-means*, sejalan dengan prinsip Islam yang mendorong pencarian pengetahuan dan kebijaksanaan. Dalam konteks pendidikan, teknologi dapat membantu guru dan institusi pendidikan dalam mengidentifikasi kebutuhan dan potensi siswa serta mengambil keputusan yang tepat dalam proses pembelajaran.

Dalam konteks Islam, penerapan teknologi harus selaras dengan nilai-nilai moral dan etika yang diajarkan oleh agama. Misalnya, penggunaan data siswa harus dilakukan dengan menghormati privasi dan hak-hak individu serta mempromosikan keadilan dan kesetaraan (Al-Farabi, 2016). Selain itu, teknologi harus digunakan untuk meningkatkan kualitas pendidikan dan memastikan bahwa setiap siswa memiliki kesempatan yang sama untuk belajar dan berkembang.

2.5 Landasan Teori

Landasan teori adalah dasar konsep dan teori yang digunakan sebagai acuan dalam penelitian ini. Landasan teori memberikan pemahaman yang lebih mendalam tentang konsep-konsep yang digunakan dalam penerapan algoritma K-Means untuk penilaian hasil belajar siswa. Bagian ini akan membahas beberapa konsep utama, termasuk clustering, algoritma K-Means, dan konsep-konsep lain yang relevan.

2.5.1 Teori Pembelajaran

Teori pembelajaran merupakan kumpulan prinsip dan konsep yang menjelaskan bagaimana siswa memperoleh, mengorganisasi, dan mempertahankan pengetahuan serta keterampilan baru (Hew dkk., 2020). Ada beberapa teori pembelajaran yang terkenal, seperti behaviorisme, kognitivisme, dan konstruktivisme. Semua teori ini memiliki tujuan yang sama, yaitu memahami bagaimana proses pembelajaran terjadi dan bagaimana guru dapat mendukung siswa dalam mencapai hasil belajar yang optimal (Slavin, 2012).

2.5.2 Teori Evaluasi Pendidikan

Teori evaluasi pendidikan merujuk pada kumpulan konsep dan metode yang digunakan untuk mengukur efektivitas dan kualitas proses pembelajaran dan pengajaran. Evaluasi dalam pendidikan melibatkan pengumpulan dan analisis informasi tentang prestasi siswa, kemajuan, dan kebutuhan, serta keberhasilan program pendidikan secara keseluruhan. Salah satu teori evaluasi pendidikan yang terkenal adalah CIPP Model (*Context, Input, Process, and Product*) yang dikembangkan oleh Daniel Stufflebeam (Hew dkk., 2020).

2.5.3 Teori Pemrosesan Data dan Analisis

Teori pemrosesan data dan analisis berkaitan dengan cara mengumpulkan, mengelola, dan menganalisis data untuk menghasilkan informasi yang berguna dan relevan. Dalam konteks pendidikan, teori ini menjadi penting karena membantu guru dan institusi pendidikan untuk mengidentifikasi pola dan tren dalam hasil belajar siswa serta membuat keputusan yang tepat untuk meningkatkan kualitas pendidikan. Salah satu metode analisis data yang sering digunakan dalam penelitian

pendidikan adalah analisis kluster, seperti algoritma *K-means*, yang memungkinkan pengelompokan siswa berdasarkan prestasi dan karakteristik tertentu (Ganga dkk., 2014).

2.6 Hubungan antara variabel penelitian

Dalam penelitian ini, terdapat beberapa variabel saling berhubungan dan mempengaruhi satu sama lain. Berikut adalah hubungan antara variabel penelitian:

- a. Variabel independen: Algoritma *K-means* merupakan teknik pengelompokan data yang digunakan dalam penelitian ini untuk mengelompokkan siswa berdasarkan prestasi belajar mereka.
- b. Variabel dependen: Penilaian hasil belajar siswa merupakan variabel yang dipengaruhi oleh penerapan algoritma *K-means*. Nilai dari mata pelajaran mewakili kemampuan akademik siswa dalam berbagai bidang. Nilai ini memberi gambaran mengenai pemahaman siswa terhadap materi yang diajarkan. Algoritma K-Means mengelompokkan siswa berdasarkan kemiripan nilai, menghasilkan kluster yang menunjukkan pola prestasi belajar. Setiap kluster mewakili kelompok siswa dengan karakteristik nilai yang serupa. Melalui pengelompokan yang efektif, sistem penilaian hasil belajar diharapkan menjadi lebih akurat dan membantu guru dalam mengidentifikasi kebutuhan dan potensi siswa. Dengan mengetahui pola prestasi dari hasil klusterisasi, sekolah dapat merancang strategi yang lebih tepat, seperti program remedial untuk kluster rendah, atau program pengayaan bagi siswa di kluster tinggi.
- c. Variabel mediasi: Integrasi nilai-nilai Islam berperan sebagai variabel mediasi yang mempengaruhi bagaimana algoritma *K-means* diterapkan dalam konteks

pendidikan Islam. Variabel ini memastikan bahwa penerapan algoritma *K-means* selaras dengan prinsip-prinsip pendidikan Islam.

Hubungan antara variabel penelitian dapat dijelaskan sebagai berikut: Penerapan algoritma *K-means* (variabel independen) mempengaruhi penilaian hasil belajar siswa (variabel dependen), di mana integrasi nilai-nilai Islam (variabel mediasi) memastikan bahwa proses tersebut sesuai dengan prinsip-prinsip pendidikan Islam. Dengan memahami hubungan antara variabel penelitian, penelitian ini dapat menghasilkan temuan yang relevan dan bermanfaat dalam meningkatkan efektivitas sistem penilaian hasil belajar siswa, terutama dalam konteks pendidikan Islam.

BAB III

DESAIN DAN IMPLEMENTASI

3.1 Data

Data yang digunakan dalam penelitian ini adalah data hasil belajar siswa. Data ini meliputi nilai-nilai siswa di kelas 3 MI Miftahul Uquiyah pada mata pelajaran matematika, bahasa Indonesia, ilmu pengetahuan alam, dan agama. Data ini akan digunakan sebagai input untuk algoritma *K-means* dalam mengelompokkan siswa berdasarkan pola prestasi belajar mereka. Dalam penelitian ini data berupa nilai siswa yaitu nilai pengetahuan dan nilai keterampilan yang masing-masing dari nilai rata-rata nilai matematika, bahasa Indonesia, ilmu pengetahuan alam dan agama.

Nilai pengetahuan merupakan nilai yang memuat pemahaman siswa kepada materi pelajaran yang diberikan. Nilai pengetahuan diperoleh berdasarkan nilai pada ulangan harian, nilai pada UTS (ujian Tengah semester) dan nilai UAS (Ujian Akhir Semester). Sedangkan nilai keterampilan merupakan nilai yang memuat pemahaman siswa dalam penerapan dari teori-teori berdasarkan pengetahuan secara praktik. Nilai keterampilan diperoleh dari tugas praktek, proyek siswa, atau kegiatan lainnya yang memerlukan aplikasi keterampilan. Jadi, dalam satu mata pelajaran terdapat dua kategori nilai. Setiap jenis nilai yaitu nilai pengetahuan dan nilai keterampilan dihitung sebagai rata-rata nilai dari masing-masing mata pelajaran.

Tabel 3. 1 Data nilai siswa

NO	NAMA	NILAI PENGETAHUAN							NILAI KETERAMPILAN						
		Quran Hadits	Akidah Akhlak	Fikih	SKI	Bahasa Arab	Tematik	Rata-rata	Quran Hadits	Akidah Akhlak	Fikih	SKI	Bahasa Arab	Tematik	Rata-rata
1	SISWA 1	98	100	100	97	95	97	97.833	95	97	98	95	98	97	96.667
2	SISWA 2	88	90	83	88	89	90	88	90	90	95	90	90	90	90.833
3	SISWA 3	88	95	92	90	90	95	91.667	90	90	88	90	90	95	90.5
4	SISWA 4	90	100	98	95	89	95	94.5	95	92	90	95	90	92	92.333
5	SISWA 5	90	100	98	90	90	92	93.333	93	92	92	93	92	92	92.333
6	SISWA 6	100	100	98	90	90	92	95	98	92	90	98	90	92	93.333
7	SISWA 7	88	95	92	90	90	95	91.667	90	90	88	90	90	95	90.5
8	SISWA 8	90	95	90	100	97	92	94	90	92	90	97	90	92	91.833
9	SISWA 9	75	75	75	75	75	75	75	75	75	75	75	75	75	75
10	SISWA 10	80	84	78	77	77	78	79	78	78	80	78	80	78	78.667
11	SISWA 11	94	95	87	90	94	95	92.5	94	92	90	95	96	92	93.167
12	SISWA 12	79	85	80	90	98	100	88.667	96	90	88	75	90	90	88.167
13	SISWA 13	79	87	78	75	78	78	79.167	75	78	75	75	75	78	76
14	SISWA 14	80	87	80	80	90	90	84.5	90	90	80	90	80	90	86.667
15	SISWA 15	100	100	90	90	94	95	94.833	98	95	95	98	95	95	96
16	SISWA 16	75	75	75	75	75	75	75	75	75	75	75	75	75	75
17	SISWA 17	79	78	75	77	75	78	77	80	78	75	80	75	78	77.667
18	SISWA 18	75	75	75	75	75	75	75	75	75	75	75	75	75	75
19	SISWA 19	100	100	100	97	94	95	97.667	98	95	98	98	98	95	97

Lanjutan tabel 3.1

NO	NAMA	NILAI PENGETAHUAN							NILAI KETERAMPILAN						
		Quran Hadits	Akidah Akhlak	Fikih	SKI	Bahasa Arab	Tematik	Rata-rata	Quran Hadits	Akidah Akhlak	Fikih	SKI	Bahasa Arab	Tematik	Rata-rata
20	SISWA 20	80	78	82	80	80	80	80	90	80	82	90	82	80	84
21	SISWA 21	85	92	87	90	88	92	89	80	92	90	80	90	92	87.333
22	SISWA 22	100	100	100	98	95	97	98.333	98	97	98	98	98	97	97.667
23	SISWA 23	85	90	88	80	88	92	87.167	90	92	92	90	92	92	91.333
24	SISWA 24	79	78	80	77	75	80	78.167	78	80	82	78	82	80	80
25	SISWA 25	80	82	85	77	80	90	82.333	90	90	88	90	88	90	89.333
26	SISWA 26	79	80	75	75	78	78	77.5	75	78	80	75	80	78	77.667
27	SISWA 27	75	75	75	75	75	75	75	75	75	75	75	75	75	75
28	SISWA 28	98	100	100	97	95	97	97.833	95	97	98	95	98	97	96.667
29	SISWA 29	80	84	78	77	77	78	79	78	78	80	78	80	78	78.667
30	SISWA 30	88	95	92	90	90	95	91.667	90	90	88	90	90	95	90.5
31	SISWA 31	85	90	88	80	88	92	87.167	90	92	92	90	92	92	91.333
32	SISWA 32	85	90	88	80	88	92	87.167	90	92	92	90	92	92	91.333
33	SISWA 33	100	100	100	98	95	97	98.333	98	97	98	98	98	97	97.667
34	SISWA 34	85	92	87	90	88	92	89	80	92	90	80	90	92	87.333
35	SISWA 35	88	95	92	90	90	95	91.667	90	90	88	90	90	95	90.5
36	SISWA 36	88	95	92	90	90	95	91.667	90	90	88	90	90	95	90.5
37	SISWA 37	80	84	78	77	77	78	79	78	78	80	78	80	78	78.667
38	SISWA 38	75	75	75	75	75	75	75	75	75	75	75	75	75	75

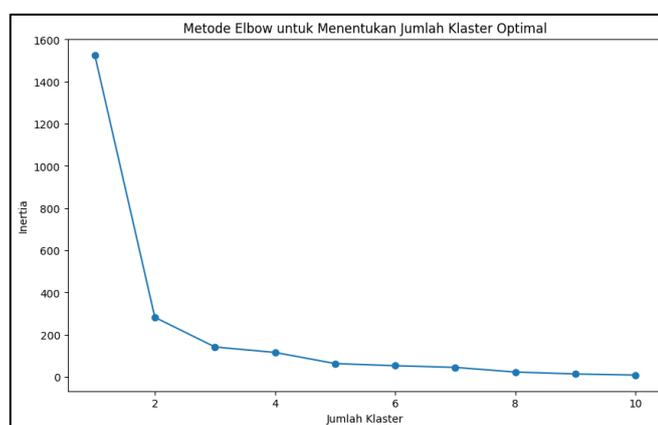
Lanjutan table 3.1

NO	NAMA	NILAI PENGETAHUAN							NILAI KETERAMPILAN						
		Quran Hadits	Akidah Akhlak	Fikih	SKI	Bahasa Arab	Tematik	Rata-rata	Quran Hadits	Akidah Akhlak	Fikih	SKI	Bahasa Arab	Tematik	Rata-rata
39	SISWA 39	88	95	92	90	90	95	91.667	90	90	88	90	90	95	90.5
40	SISWA 40	79	80	75	75	78	78	77.5	75	78	80	75	80	78	77.667
41	SISWA 41	100	100	90	90	94	95	94.833	98	95	95	98	95	95	96
42	SISWA 42	80	78	82	80	80	80	80	90	80	82	90	82	80	84
43	SISWA 43	79	80	75	75	78	78	77.5	75	78	80	75	80	78	77.667
44	SISWA 44	80	78	82	80	80	80	80	90	80	82	90	82	80	84
45	SISWA 45	79	80	75	75	78	78	77.5	75	78	80	75	80	78	77.667
46	SISWA 46	80	78	82	80	80	80	80	90	80	82	90	82	80	84
47	SISWA 47	88	95	92	90	90	95	91.667	90	90	88	90	90	95	90.5
48	SISWA 48	85	90	88	80	88	92	87.167	90	92	92	90	92	92	91.333
49	SISWA 49	94	95	87	90	94	95	92.5	94	92	90	95	96	92	93.167
50	SISWA 50	85	90	88	80	88	92	87.167	90	92	92	90	92	92	91.333
51	SISWA 51	94	95	87	90	94	95	92.5	94	92	90	95	96	92	93.167
52	SISWA 52	88	95	92	90	90	95	91.667	90	90	88	90	90	95	90.5
53	SISWA 53	100	100	90	90	94	95	94.833	98	95	95	98	95	95	96
54	SISWA 54	100	100	90	90	94	95	94.833	98	95	95	98	95	95	96

Informasi ini diperoleh melalui observasi dan wawancara dengan guru dan pihak sekolah yang relevan. Informasi ini mencakup metode pembelajaran yang digunakan serta kebijakan dan praktik penilaian hasil belajar siswa di sekolah. Penilaian ini mencakup aspek kognitif (pengetahuan) dan psikomotorik (keterampilan). Sistem penilaian ini cenderung menggunakan pendekatan yang sama untuk semua siswa, tanpa mempertimbangkan pola nilai antar mata pelajaran atau antar siswa.

3.2 Proses K-Means Clustering

Proses algoritma *K-Means* diawali dengan menentukan jumlah *cluster* dan menentukan *centroid* awal yang di tentukan secara acak. Jumlah cluster ditentukan dengan menggunakan *elbow method*. Hasil dari *elbow method* adalah grafik seperti pada gambar 3.1 berikut:



Gambar 3. 1 Grafik elbow method

Berdasarkan hasil grafik pada gambar 3.1 titik penurunan yang mulai lambat ada pada titik 3 sehingga jumlah *cluster* yang digunakan pada data sebanyak tiga

cluster yaitu pemahaman tinggi, pemahaman cukup, pemahan kurang. Parameter data disajikan pada tabel 3.2 berikut:

Tabel 3. 2 Parameter data

Jumlah Cluster	Definisi	Atribut	Jumlah Siswa
3	Pemahaman tinggi, pemahaman cukup, pemahan kurang	12	54

Setelah di dapat jumlah *cluster* yang ditentukan, tahap selanjutnya yaitu menghitung jarak dari data ke-*i* ke *centroid* ke-*c* menggunakan persamaan seperti pada rumus *Euclidean*. Jarak Euclidean antara dua titik adalah jumlah kuadrat selisih setiap koordinat dua titik. Dalam penelitian ini, selisih yang dimaksud adalah selisih nilai dari dua siswa. Perhitungan yang digunakan *x* adalah nilai siswa A dan *y* adalah nilai siswa B untuk centroid pertama, siswa C untuk centroid kedua dan siswa D untuk centroid ketiga. Berikut perhitungan Algoritma K-Means dalam menentukan jarak berdasarkan data nilai siswa :

Pada jarak ke centroid 1

$$\begin{aligned}
 d(x_i v_c) &= x_i - v_c = \sqrt{\sum_{j=1}^m (x_{ij} - v_{cj})^2} \\
 &= \sqrt{(97.833 - 88)^2 + (96.667 - 90.833)^2} \\
 &= 11.43338192
 \end{aligned}$$

Pada jarak ke centroid 2

$$\begin{aligned}
 d(x_i v_c) = x_i - v_c &= \sqrt{\sum_{j=1}^m (x_{ij} - v_{cj})^2} \\
 &= \sqrt{(97.833 - 91.667)^2 + (96.667 - 90.5)^2} \\
 &= 8.720983635
 \end{aligned}$$

Pada jarak ke centroid 3

$$\begin{aligned}
 d(x_i v_c) = x_i - v_c &= \sqrt{\sum_{j=1}^m (x_{ij} - v_{cj})^2} \\
 &= \sqrt{(97.833 - 94.5)^2 + (96.667 - 92.334)^2} \\
 &= 5.467073156
 \end{aligned}$$

Setelah perhitungan jarak selesai, masing-masing data memiliki tiga nilai jarak. Setelah data dihitung dan mendapatkan hasil pada jarak ke centroid 1, jarak ke centroid 2 dan jarak ke centroid 3. Jarak ke centroid 1 merupakan nilai jarak data ke centroid kluster 1, jarak ke centroid 2 merupakan nilai jarak data ke centroid kluster 2, jarak ke centroid 3 merupakan nilai jarak data ke centroid kluster 3. kemudian setiap hasil yang memiliki jarak terdekat yang termasuk dalam cluster tersebut. Dengan pengelompokan hasil ini siswa yang termasuk dalam satu kluster dapat dianggap memiliki pola nilai yang mirip karena jarak nilai lebih dekat ke centroid yang sama. Hasil dari perhitungan manual yang dilakukan dengan Algoritma K-Means pada data nilai siswa kelas 3 MI Miftahul Uquliyah adalah :

Tabel 3. 3 Hasil perhitungan manual Algoritma K-Means

NO	NAMA	NILAI PENGETAHUAN	NILAI KETERAMPILAN	C1	C2	C3	CLUSTER
1	SISWA 1	97.83333333	96.66666667	11.43338	8.720984	5.467073	C3
2	SISWA 2	88	90.83333333	0	3.681787	6.670832	C1
3	SISWA 3	91.66666667	90.5	3.681787	0	3.374743	C2
4	SISWA 4	94.5	92.33333333	6.670832	3.374743	0	C3
5	SISWA 5	93.33333333	92.33333333	5.540257	2.477678	1.166667	C3
6	SISWA 6	95	93.33333333	7.433034	4.374802	1.118034	C3
7	SISWA 7	91.27777778	91.7037037	3.391367	1.264965	3.283162	C2
8	SISWA 8	94	91.83333333	6.082763	2.687419	0.707107	C3
9	SISWA 9	75	75	20.48645	22.76022	26.09012	C1
10	SISWA 10	79	78.66666667	15.13366	17.33413	20.66465	C1
11	SISWA 11	92.5	93.16666667	5.068969	2.793842	2.166667	C3
12	SISWA 12	88.66666667	88.16666667	2.748737	3.800585	7.168604	C1
13	SISWA 13	79.16666667	76	17.26429	19.14419	22.40288	C1
14	SISWA 14	84.5	86.66666667	5.441609	8.127457	11.49396	C1
15	SISWA 15	94.83333333	96	8.566732	6.346478	3.681787	C3
16	SISWA 16	78.5	77.66666667	16.23611	18.38629	21.70509	C1
17	SISWA 17	77	77.66666667	17.15696	19.4886	22.83333	C1
18	SISWA 18	75	75	20.48645	22.76022	26.09012	C1
19	SISWA 19	97.66666667	97	11.46613	8.845903	5.639641	C3
20	SISWA 20	80	84	10.52114	13.35519	16.72407	C1

Lanjutan tabel 3.3

NO	NAMA	NILAI PENGETAHUAN	NILAI KETERAMPILAN	C1	C2	C3	CLUSTER
21	SISWA 21	89	87.33333333	3.640055	4.139914	7.433034	C1
22	SISWA 22	98.33333333	97.66666667	12.38839	9.788031	6.56802	C2
23	SISWA 23	87.16666667	91.33333333	0.971825	4.57651	7.401201	C1
24	SISWA 24	78.16666667	80	14.63064	17.10263	20.46678	C1
25	SISWA 25	82.33333333	89.33333333	5.861835	9.405967	12.53107	C1
26	SISWA 26	77.5	77.66666667	16.84076	19.11515	22.45242	C1
27	SISWA 27	75	75	20.48645	22.76022	26.09012	C2
28	SISWA 28	97.83333333	96.66666667	11.43338	8.720984	5.467073	C3
29	SISWA 29	79	78.66666667	15.13366	17.33413	20.66465	C4
30	SISWA 30	91.66666667	90.5	3.681787	0	3.374743	C5
31	SISWA 31	87.16666667	91.33333333	0.971825	4.57651	7.401201	C6
32	SISWA 32	87.16666667	91.33333333	0.971825	4.57651	7.401201	C7
33	SISWA 33	98.33333333	97.66666667	12.38839	9.788031	6.56802	C8
34	SISWA 34	89	87.33333333	3.640055	4.139914	7.433034	C9
35	SISWA 35	91.66666667	90.5	3.681787	0	3.374743	C10
36	SISWA 36	91.66666667	90.5	3.681787	0	3.374743	C11
37	SISWA 37	79	78.66666667	15.13366	17.33413	20.66465	C12
38	SISWA 38	75	75	20.48645	22.76022	26.09012	C13
39	SISWA 39	91.66666667	90.5	3.681787	0	3.374743	C14
40	SISWA 40	77.5	77.66666667	16.84076	19.11515	22.45242	C15

Lanjutan tabel 3.3

NO	NAMA	NILAI PENGETAHUAN	NILAI KETERAMPILAN	C1	C2	C3	CLUSTER
41	SISWA 41	94.83333333	96	8.566732	6.346478	3.681787	C16
42	SISWA 42	80	84	10.52114	13.35519	16.72407	C17
43	SISWA 43	77.5	77.66666667	16.84076	19.11515	22.45242	C18
44	SISWA 44	80	84	10.52114	13.35519	16.72407	C19
45	SISWA 45	77.5	77.66666667	16.84076	19.11515	22.45242	C20
46	SISWA 46	80	84	10.52114	13.35519	16.72407	C21
47	SISWA 47	91.66666667	90.5	3.681787	0	3.374743	C22
48	SISWA 48	87.16666667	91.33333333	0.971825	4.57651	7.401201	C23
49	SISWA 49	92.5	93.16666667	5.068969	2.793842	2.166667	C24
50	SISWA 50	87.16666667	91.33333333	0.971825	4.57651	7.401201	C25
51	SISWA 51	92.5	93.16666667	5.068969	2.793842	2.166667	C26
52	SISWA 52	91.66666667	90.5	3.681787	0	3.374743	C27
53	SISWA 53	94.83333333	96	8.566732	6.346478	3.681787	C28
54	SISWA 54	94.83333333	96	8.566732	6.346478	3.681787	C29

Langkah selanjutnya yaitu menghitung ulang centroid untuk setiap cluster menggunakan persamaan (2.2). Jika terdapat perubahan yang signifikan antara centroid yang baru dengan centroid yang lama maka akan diulang pada proses perhitungan jarak, tetapi jika tidak maka selesai.

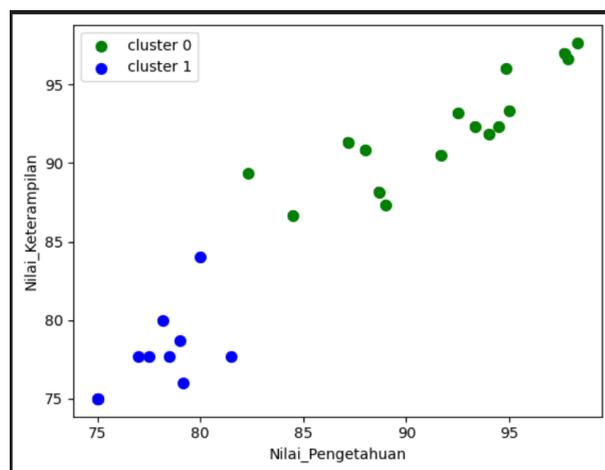
- i. Uji Coba dan Evaluasi: Menguji coba sistem yang telah dikembangkan pada kelompok siswa yang dipilih untuk mengevaluasi kinerja sistem dalam konteks nyata. Mengumpulkan data hasil uji coba, seperti nilai siswa, waktu yang dibutuhkan untuk menyelesaikan tes, dan kepuasan siswa terhadap sistem. Melakukan analisis statistik terhadap data hasil uji coba untuk mengevaluasi kinerja sistem.
- ii. Analisis dan Interpretasi Hasil: Menganalisis hasil evaluasi sistem, termasuk perbandingan antara metode penilaian tradisional dan sistem yang dikembangkan. Menginterpretasikan hasil analisis dalam konteks penelitian, serta membahas implikasi dan dampak dari sistem yang dikembangkan.
- iii. Revisi dan Penyempurnaan Sistem: Melakukan revisi dan penyempurnaan sistem berdasarkan hasil analisis dan interpretasi, serta umpan balik dari para pengguna. Mengembangkan versi akhir sistem penilaian hasil belajar siswa yang efektif, efisien, dan sesuai dengan nilai-nilai Islam.
- iv. Penyusunan Laporan Penelitian: Menyusun laporan penelitian yang mencakup metodologi yang digunakan, hasil pengembangan sistem, analisis dan interpretasi hasil, serta kesimpulan dan rekomendasi untuk penelitian selanjutnya.

3.3 Implementasi

Implementasi merupakan tahapan memasukkan rumus ke dalam program. Peneliti menggunakan Bahasa pemrograman *python* untuk menentukan *cluster* siswa. *Scikit-learn* merupakan module dari bahasa pemrograman *python* yang menyediakan fungsi *K-Means*. Data ditampilkan dalam sebuah *scatterplot* menggunakan *matplotlib*. *Matplotlib* merupakan *library* dalam Bahasa *python* untuk menciptakan visualisasi data. Berikut merupakan *source code* dari proses *clustering*:

1. `plt.scatter(x[:,0],x[:,1],c=kmeans.labels_,marker='o')`
2. `centroids = kmeans.cluster_centers_`
3. `labels = kmeans.labels_`
4. `plt.scatter(x[labels ==0,0], x[labels == 0,1], s=50, color='green',label='cluster 1')`
5. `plt.scatter(x[labels ==1,0], x[labels == 1,1], s=50, color='blue',label='cluster 2')`
- 6.
7. `plt.legend()`
8. `plt.xlabel('Nilai_Pengetahuan')`
9. `plt.ylabel('Nilai_Keterampilan')`
10. `plt.show()`

Berdasarkan source code diatas maka hasilnya seperti gambar berikut:



Gambar 3. 2 Scatterplot

Untuk menguji algoritma *K-Means*, pengujian sensitivitas terhadap jumlah cluster merupakan salah satu langkah penting saat menggunakan algoritma *K-Means*. Tujuan dari pengujian ini adalah untuk menentukan jumlah *cluster* atau *cluster* yang optimal yang paling sesuai menggunakan metode *Elbow*. Source code untuk menampilkan hasil metode *Elbow* sebagai berikut:

```
1. wcss = []
2. for i in range(1, 11):
3.     kmeans = KMeans(n_clusters=i, init='k-means++', max_iter=30, n_init=2, random_state=2)
4.     kmeans.fit(x)
5.     wcss.append(kmeans.inertia_)
6.
7. plt.plot(range(1, 11), wcss)
8. plt.title('Elbow Method')
9. plt.xlabel('Number of clusters')
10. plt.ylabel('Within-cluster sum of squares')
11. plt.show()
```

Hasil yang ditampilkan merupakan sebuah grafik pada gambar 3.1 yang terdapat titik yang tampak seperti siku tangan manusia. Titik yang dicari tersebut disebut dengan *elbow point*

BAB IV

HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil pengujian dan analisis yang telah dilakukan dalam penelitian. Termasuk proses, hasil, evaluasi, serta interpretasi dari pengujian yang menggunakan algoritma k-means yang dimulai dari pengolahan data mentah yaitu nilai pada mata pelajaran yang di bedakan oleh nilai pengetahuan dan nilai keterampilan.

4.1 Hasil Penelitian

Dalam penelitian ini, hasil penelitian digunakan untuk menjawab rumusan masalah yang telah di tuliskan pada BAB I. Untuk itu perlu dilakukan pengujian terhadap kinerja algoritma *K-Means clustering* dalam melakukan pengelompokan siswa. Pengujian dilakukan dengan menggunakan data sebanyak 26 data nilai siswa dengan 5 skenario percobaan

Tabel 4. 1 Data nilai siswa

No	Nama	Nilai Pengetahuan	Nilai Keterampilan
1	SISWA 1	97.83333333	96.66666667
2	SISWA 2	88	90.83333333
3	SISWA 3	91.66666667	90.5
4	SISWA 4	94.5	92.33333333
5	SISWA 5	93.33333333	92.33333333
6	SISWA 6	95	93.33333333
7	SISWA 7	91.27777778	91.7037037
8	SISWA 8	94	91.83333333
9	SISWA 9	75	75
10	SISWA 10	79	78.66666667
11	SISWA 11	92.5	93.16666667
12	SISWA 12	88.66666667	88.16666667

Lanjutan tabel 4.1

No	Nama	Nilai Pengetahuan	Nilai Keterampilan
13	SISWA 13	79.16666667	76
14	SISWA 14	84.5	86.66666667
15	SISWA 15	94.83333333	96
16	SISWA 16	78.5	77.66666667
17	SISWA 17	77	77.66666667
18	SISWA 18	75	75
19	SISWA 19	97.66666667	97
20	SISWA 20	80	84
21	SISWA 21	89	87.33333333
22	SISWA 22	98.33333333	97.66666667
23	SISWA 23	87.16666667	91.33333333
24	SISWA 24	78.16666667	80
25	SISWA 25	82.33333333	89.33333333
26	SISWA 26	77.5	77.66666667
27	SISWA 27	75	75
28	SISWA 28	97.83333333	96.66666667
29	SISWA 29	79	78.66666667
30	SISWA 30	91.66666667	90.5
31	SISWA 31	87.16666667	91.33333333
32	SISWA 32	87.16666667	91.33333333
33	SISWA 33	98.33333333	97.66666667
34	SISWA 34	89	87.33333333
35	SISWA 35	91.66666667	90.5
36	SISWA 36	91.66666667	90.5
37	SISWA 37	79	78.66666667
38	SISWA 38	75	75
39	SISWA 39	91.66666667	90.5
40	SISWA 40	77.5	77.66666667
41	SISWA 41	94.83333333	96
42	SISWA 42	80	84
43	SISWA 43	77.5	77.66666667
44	SISWA 44	80	84
45	SISWA 45	77.5	77.66666667
46	SISWA 46	80	84
47	SISWA 47	91.66666667	90.5

Lanjutan tabel 4.1

No	Nama	Nilai Pengetahuan	Nilai Keterampilan
48	SISWA 48	87.16666667	91.33333333
49	SISWA 49	92.5	93.16666667
50	SISWA 50	87.16666667	91.33333333
51	SISWA 51	92.5	93.16666667
52	SISWA 52	91.66666667	90.5
53	SISWA 53	94.83333333	96
54	SISWA 54	94.83333333	96

Pengujian yang dilakukan adalah dengan 5 kali pengujian menggunakan centroid awal yang berbeda. Hasil dari pengujian menghasilkan data yang berbeda. Selain itu dari proses pengujian juga dihitung silhouette coefficient masing-masing.

Tabel 4. 2 Data uji

Pengujian ke-	Data Uji
1	2, 3, 4
2	26, 7, 15
3	9, 21, 22
4	17, 12, 6
5	11, 16, 19

4.1.1 Uji Coba ke-1

Pengujian ke-1 menggunakan scenario data uji 2,3,4 atau nilai centroid 1, nilai centroid 2 dan nilai centroid 3. Nilai centroid 1 diambil dari siswa baris ke-2, nilai centroid 2 diambil dari siswa baris ke 3, dan nilai centroid 3 diambil dari siswa baris ke-4. Pada pengujian ini diperoleh hasil berikut ini:

Tabel 4. 3 Hasil uji coba 1

No	NAMA	Jarak ke centroid 1	Jarak ke centroid 2	Jarak ke centroid 3	cluster
1	SISWA 1	13.97446	5.605116	1.29704058	3
2	SISWA 2	7.787617	0.890782	5.049092978	2
3	SISWA 3	9.534374	1.322897	3.175656292	2

Lanjutan tabel 4.3

No	NAMA	Jarak ke centroid 1	Jarak ke centroid 2	Jarak ke centroid 3	cluster
4	SISWA 4	11.37574	3.115481	1.455055347	3
5	SISWA 5	10.76762	2.442212	1.93616232	3
6	SISWA 6	11.83314	3.517597	0.955956447	3
7	SISWA 7	9.534374	1.322897	3.175656292	2
8	SISWA 8	11.01654	2.778626	1.79400895	3
9	SISWA 9	2.2602	10.65235	14.92881135	1
10	SISWA 10	0.668568	7.849092	12.11750192	1
11	SISWA 11	10.52085	2.122691	2.195860068	2
12	SISWA 12	7.50915	1.029444	5.183603971	2
13	SISWA 13	1.303267	8.458906	12.65668661	1
14	SISWA 14	5.1067	3.339261	7.648714079	2
15	SISWA 15	12.31514	3.914046	0.66865581	3
16	SISWA 16	0.533576	8.330324	12.58930178	1
17	SISWA 17	0.656921	9.057781	13.34347075	1
18	SISWA 18	2.2602	10.65235	14.92881135	1
19	SISWA 19	13.95863	5.577096	1.270442315	3
20	SISWA 20	2.494009	6.235509	10.54230418	1
21	SISWA 21	7.522175	1.275946	5.216247439	2
22	SISWA 22	14.4391	6.055681	1.750899083	3
23	SISWA 23	7.53796	1.422175	5.456800639	2
24	SISWA 24	0.572693	7.962016	12.26242637	1
25	SISWA 25	5.044292	4.300406	8.464284326	2
26	SISWA 26	0.450109	8.811755	13.09003365	1
27	SISWA 27	2.2602	10.65235	14.92881135	1
28	SISWA 28	13.97446	5.605116	1.29704058	3
29	SISWA 29	0.668568	7.849092	12.11750192	1
30	SISWA 30	9.534374	1.322897	3.175656292	2
31	SISWA 31	7.53796	1.422175	5.456800639	2
32	SISWA 32	7.53796	1.422175	5.456800639	2
33	SISWA 33	14.4391	6.055681	1.750899083	3
34	SISWA 34	7.522175	1.275946	5.216247439	2
35	SISWA 35	9.534374	1.322897	3.175656292	2
36	SISWA 36	9.534374	1.322897	3.175656292	2
37	SISWA 37	0.668568	7.849092	12.11750192	1

Lanjutan tabel 4.3

No	NAMA	Jarak ke centroid 1	Jarak ke centroid 2	Jarak ke centroid 3	cluster
38	SISWA 38	2.2602	10.65235	14.92881135	1
39	SISWA 39	9.534374	1.322897	3.175656292	2
40	SISWA 40	0.450109	8.811755	13.09003365	1
41	SISWA 41	12.31514	3.914046	0.66865581	3
42	SISWA 42	2.494009	6.235509	10.54230418	1
43	SISWA 43	0.450109	8.811755	13.09003365	1
44	SISWA 44	2.494009	6.235509	10.54230418	1
45	SISWA 45	0.450109	8.811755	13.09003365	1
46	SISWA 46	2.494009	6.235509	10.54230418	1
47	SISWA 47	9.534374	1.322897	3.175656292	2
48	SISWA 48	7.53796	1.422175	5.456800639	2
49	SISWA 49	10.52085	2.122691	2.195860068	2
50	SISWA 50	7.53796	1.422175	5.456800639	2
51	SISWA 51	10.52085	2.122691	2.195860068	2
52	SISWA 52	9.534374	1.322897	3.175656292	2
53	SISWA 53	12.31514	3.914046	0.66865581	3
54	SISWA 54	12.31514	3.914046	0.66865581	3

Dari hasil percobaan diatas dihasilkan kelompok siswa dengan pemahaman tinggi sebanyak 13 siswa, pemahaman sedang sebanyak 22 siswa, dan pemahaman kurang sebanyak 19 siswa.

4.1.2 Uji Coba ke-2

Pengujian ke-2 menggunakan skenario data uji 26,7,15 atau nilai centroid 1, nilai centroid 2 dan nilai centroid 3. Nilai centroid 1 diambil dari baris ke-26, nilai centroid 2 diambil dari baris ke 7, dan nilai centroid 3 diambil dari baris ke-15. Pada pengujian ini diperoleh hasil berikut ini:

Tabel 4. 4 Hasil uji coba 2

No	NAMA	Jarak ke centroid 1	Jarak ke centroid 2	Jarak ke centroid 3	cluster
1	SISWA 1	13.97446	5.605116	1.29704058	3
2	SISWA 2	7.787617	0.890782	5.049092978	2
3	SISWA 3	9.534374	1.322897	3.175656292	2
4	SISWA 4	11.37574	3.115481	1.455055347	3
5	SISWA 5	10.76762	2.442212	1.93616232	3
6	SISWA 6	11.83314	3.517597	0.955956447	3
7	SISWA 7	9.534374	1.322897	3.175656292	2
8	SISWA 8	11.01654	2.778626	1.79400895	3
9	SISWA 9	2.2602	10.65235	14.92881135	1
10	SISWA 10	0.668568	7.849092	12.11750192	1
11	SISWA 11	10.52085	2.122691	2.195860068	2
12	SISWA 12	7.50915	1.029444	5.183603971	2
13	SISWA 13	1.303267	8.458906	12.65668661	1
14	SISWA 14	5.1067	3.339261	7.648714079	2
15	SISWA 15	12.31514	3.914046	0.66865581	3
16	SISWA 16	0.533576	8.330324	12.58930178	1
17	SISWA 17	0.656921	9.057781	13.34347075	1
18	SISWA 18	2.2602	10.65235	14.92881135	1
19	SISWA 19	13.95863	5.577096	1.270442315	3
20	SISWA 20	2.494009	6.235509	10.54230418	1
21	SISWA 21	7.522175	1.275946	5.216247439	2
22	SISWA 22	14.4391	6.055681	1.750899083	3
23	SISWA 23	7.53796	1.422175	5.456800639	2
24	SISWA 24	0.572693	7.962016	12.26242637	1
25	SISWA 25	5.044292	4.300406	8.464284326	2
26	SISWA 26	0.450109	8.811755	13.09003365	1
27	SISWA 27	2.2602	10.65235	14.92881135	1
28	SISWA 28	13.97446	5.605116	1.29704058	3
29	SISWA 29	0.668568	7.849092	12.11750192	1
30	SISWA 30	9.534374	1.322897	3.175656292	2
31	SISWA 31	7.53796	1.422175	5.456800639	2
32	SISWA 32	7.53796	1.422175	5.456800639	2
33	SISWA 33	14.4391	6.055681	1.750899083	3
34	SISWA 34	7.522175	1.275946	5.216247439	2
35	SISWA 35	9.534374	1.322897	3.175656292	2
36	SISWA 36	9.534374	1.322897	3.175656292	2
37	SISWA 37	0.668568	7.849092	12.11750192	1
38	SISWA 38	2.2602	10.65235	14.92881135	1
39	SISWA 39	9.534374	1.322897	3.175656292	2
40	SISWA 40	0.450109	8.811755	13.09003365	1

Lanjutan tabel 4.4

No	NAMA	Jarak ke centroid 1	Jarak ke centroid 2	Jarak ke centroid 3	cluster
41	SISWA 41	12.31514	3.914046	0.66865581	3
42	SISWA 42	2.494009	6.235509	10.54230418	1
43	SISWA 43	0.450109	8.811755	13.09003365	1
44	SISWA 44	2.494009	6.235509	10.54230418	1
45	SISWA 45	0.450109	8.811755	13.09003365	1
46	SISWA 46	2.494009	6.235509	10.54230418	1
47	SISWA 47	9.534374	1.322897	3.175656292	2
48	SISWA 48	7.53796	1.422175	5.456800639	2
49	SISWA 49	10.52085	2.122691	2.195860068	2
50	SISWA 50	7.53796	1.422175	5.456800639	2
51	SISWA 51	10.52085	2.122691	2.195860068	2
52	SISWA 52	9.534374	1.322897	3.175656292	2
53	SISWA 53	12.31514	3.914046	0.66865581	3
54	SISWA 54	12.31514	3.914046	0.66865581	3

Dari hasil percobaan diatas dihasilkan kelompok siswa dengan pemahaman tinggi sebanyak 13 siswa, pemahaman sedang sebanyak 22 siswa, dan pemahaman kurang sebanyak 19 siswa.

4.1.3 Uji Coba ke-3

Pengujian ke-3 menggunakan scenario data uji 9,21,22 atau nilai centroid 1, nilai centroid 2 dan nilai centroid 3. Nilai centroid 1 diambil dari baris ke-9, nilai centroid 2 diambil dari baris ke 21, dan nilai centroid 3 diambil dari baris ke-22.

Pada pengujian ini diperoleh hasil berikut ini:

Tabel 4. 5 Hasil uji coba 3

No	NAMA	Jarak ke centroid 1	Jarak ke centroid 2	Jarak ke centroid 3	cluster
1	SISWA 1	13.97446	5.605116	1.29704058	3
2	SISWA 2	7.787617	0.890782	5.049092978	2
3	SISWA 3	9.534374	1.322897	3.175656292	2
4	SISWA 4	11.37574	3.115481	1.455055347	3
5	SISWA 5	10.76762	2.442212	1.93616232	3

Lanjutan tabel 4.5

No	NAMA	Jarak ke centroid 1	Jarak ke centroid 2	Jarak ke centroid 3	cluster
6	SISWA 6	11.83314	3.517597	0.955956447	3
7	SISWA 7	9.534374	1.322897	3.175656292	2
8	SISWA 8	11.01654	2.778626	1.79400895	3
9	SISWA 9	2.2602	10.65235	14.92881135	1
10	SISWA 10	0.668568	7.849092	12.11750192	1
11	SISWA 11	10.52085	2.122691	2.195860068	2
12	SISWA 12	7.50915	1.029444	5.183603971	2
13	SISWA 13	1.303267	8.458906	12.65668661	1
14	SISWA 14	5.1067	3.339261	7.648714079	2
15	SISWA 15	12.31514	3.914046	0.66865581	3
16	SISWA 16	0.533576	8.330324	12.58930178	1
17	SISWA 17	0.656921	9.057781	13.34347075	1
18	SISWA 18	2.2602	10.65235	14.92881135	1
19	SISWA 19	13.95863	5.577096	1.270442315	3
20	SISWA 20	2.494009	6.235509	10.54230418	1
21	SISWA 21	7.522175	1.275946	5.216247439	2
22	SISWA 22	14.4391	6.055681	1.750899083	3
23	SISWA 23	7.53796	1.422175	5.456800639	2
24	SISWA 24	0.572693	7.962016	12.26242637	1
25	SISWA 25	5.044292	4.300406	8.464284326	2
26	SISWA 26	0.450109	8.811755	13.09003365	1
27	SISWA 27	2.2602	10.65235	14.92881135	1
28	SISWA 28	13.97446	5.605116	1.29704058	3
29	SISWA 29	0.668568	7.849092	12.11750192	1
30	SISWA 30	9.534374	1.322897	3.175656292	2
31	SISWA 31	7.53796	1.422175	5.456800639	2
32	SISWA 32	7.53796	1.422175	5.456800639	2
33	SISWA 33	14.4391	6.055681	1.750899083	3
34	SISWA 34	7.522175	1.275946	5.216247439	2
35	SISWA 35	9.534374	1.322897	3.175656292	2
36	SISWA 36	9.534374	1.322897	3.175656292	2
37	SISWA 37	0.668568	7.849092	12.11750192	1
38	SISWA 38	2.2602	10.65235	14.92881135	1
39	SISWA 39	9.534374	1.322897	3.175656292	2

Lanjutan tabel 4.5

No	NAMA	Jarak ke centroid 1	Jarak ke centroid 2	Jarak ke centroid 3	cluster
40	SISWA 40	0.450109	8.811755	13.09003365	1
41	SISWA 41	12.31514	3.914046	0.66865581	3
42	SISWA 42	2.494009	6.235509	10.54230418	1
43	SISWA 43	0.450109	8.811755	13.09003365	1
44	SISWA 44	2.494009	6.235509	10.54230418	1
45	SISWA 45	0.450109	8.811755	13.09003365	1
46	SISWA 46	2.494009	6.235509	10.54230418	1
47	SISWA 47	9.534374	1.322897	3.175656292	2
48	SISWA 48	7.53796	1.422175	5.456800639	2
49	SISWA 49	10.52085	2.122691	2.195860068	2
50	SISWA 50	7.53796	1.422175	5.456800639	2
51	SISWA 51	10.52085	2.122691	2.195860068	2
52	SISWA 52	9.534374	1.322897	3.175656292	2
53	SISWA 53	12.31514	3.914046	0.66865581	3
54	SISWA 54	12.31514	3.914046	0.66865581	3

Dari hasil percobaan diatas dihasilkan kelompok siswa dengan pemahaman tinggi sebanyak 13 siswa, pemahaman sedang sebanyak 22 siswa, dan pemahaman kurang sebanyak 19 siswa.

4.1.4 Uji Coba ke-4

Pengujian ke-4 menggunakan scenario data uji 17,12,6 atau nilai centroid 1, nilai centroid 2 dan nilai centroid 3. Nilai centroid 1 diambil dari baris ke-17, nilai centroid 2 diambil dari baris ke 12, dan nilai centroid 3 diambil dari baris ke-6. Pada pengujian ini diperoleh hasil berikut ini:

Tabel 4. 6 Hasil uji coba 4

No	NAMA	Jarak ke centroid 1	Jarak ke centroid 2	Jarak ke centroid 3	cluster
1	SISWA 1	13.97446	5.939312	1.708744869	3
2	SISWA 2	7.787617	0.673766	4.63777702	2

Lanjutan tabel 4.6

No	NAMA	Jarak ke centroid 1	Jarak ke centroid 2	Jarak ke centroid 3	cluster
3	SISWA 3	9.534374	1.622047	2.778404441	2
4	SISWA 4	11.37574	3.43768	1.12981224	3
5	SISWA 5	10.76762	2.770002	1.542392849	3
6	SISWA 6	11.83314	3.846429	0.656233467	3
7	SISWA 7	9.534374	1.622047	2.778404441	2
8	SISWA 8	11.01654	3.096902	1.444175323	3
9	SISWA 9	2.2602	10.31831	14.52306893	1
10	SISWA 10	0.668568	7.515889	11.71227173	1
11	SISWA 11	10.52085	2.457852	1.784136305	3
12	SISWA 12	7.50915	0.764212	4.779014089	2
13	SISWA 13	1.303267	8.1317	12.25814143	1
14	SISWA 14	5.1067	3.004964	7.237850211	2
15	SISWA 15	12.31514	4.248548	0.486858553	3
16	SISWA 16	0.533576	7.997784	12.18509235	1
17	SISWA 17	0.656921	8.723279	12.93653646	1
18	SISWA 18	2.2602	10.31831	14.52306893	1
19	SISWA 19	13.95863	5.911881	1.679145991	3
20	SISWA 20	2.494009	5.902555	10.13087021	1
21	SISWA 21	7.522175	1.073736	4.818780174	2
22	SISWA 22	14.4391	6.390576	2.158852093	3
23	SISWA 23	7.53796	1.209943	5.047752867	2
24	SISWA 24	0.572693	7.62689	11.85356618	1
25	SISWA 25	5.044292	3.99885	8.055679718	2
26	SISWA 26	0.450109	8.477731	12.68390617	1
27	SISWA 27	2.2602	10.31831	14.52306893	1
28	SISWA 28	13.97446	5.939312	1.708744869	3
29	SISWA 29	0.668568	7.515889	11.71227173	1
30	SISWA 30	9.534374	1.622047	2.778404441	2
31	SISWA 31	7.53796	1.209943	5.047752867	2
32	SISWA 32	7.53796	1.209943	5.047752867	2
33	SISWA 33	14.4391	6.390576	2.158852093	3
34	SISWA 34	7.522175	1.073736	4.818780174	2
35	SISWA 35	9.534374	1.622047	2.778404441	2
36	SISWA 36	9.534374	1.622047	2.778404441	2

Lanjutan tabel 4.6

No	NAMA	Jarak ke centroid 1	Jarak ke centroid 2	Jarak ke centroid 3	cluster
37	SISWA 37	0.668568	7.515889	11.71227173	1
38	SISWA 38	2.2602	10.31831	14.52306893	1
39	SISWA 39	9.534374	1.622047	2.778404441	2
40	SISWA 40	0.450109	8.477731	12.68390617	1
41	SISWA 41	12.31514	4.248548	0.486858553	3
42	SISWA 42	2.494009	5.902555	10.13087021	1
43	SISWA 43	0.450109	8.477731	12.68390617	1
44	SISWA 44	2.494009	5.902555	10.13087021	1
45	SISWA 45	0.450109	8.477731	12.68390617	1
46	SISWA 46	2.494009	5.902555	10.13087021	1
47	SISWA 47	9.534374	1.622047	2.778404441	2
48	SISWA 48	7.53796	1.209943	5.047752867	2
49	SISWA 49	10.52085	2.457852	1.784136305	3
50	SISWA 50	7.53796	1.209943	5.047752867	2
51	SISWA 51	10.52085	2.457852	1.784136305	3
52	SISWA 52	9.534374	1.622047	2.778404441	2
53	SISWA 53	12.31514	4.248548	0.486858553	3
54	SISWA 54	12.31514	4.248548	0.486858553	3

Dari hasil percobaan diatas dihasilkan kelompok siswa dengan pemahaman tinggi sebanyak 16 siswa, pemahaman sedang sebanyak 19 siswa, dan pemahaman kurang sebanyak 19 siswa.

4.1.5 Uji Coba ke-5

Pengujian ke-5 menggunakan scenario data uji 11,16,19 atau nilai centroid 1, nilai centroid 2 dan nilai centroid 3. Nilai centroid 1 diambil dari baris ke-11, nilai centroid 2 diambil dari baris ke-16, dan nilai centroid 3 diambil dari baris ke-19 Pada pengujian ini diperoleh hasil berikut ini:

Tabel 4. 7 Hasil uji coba 5

No	NAMA	Jarak ke centroid 1	Jarak ke centroid 2	Jarak ke centroid 3	cluster
1	SISWA 1	5.605116	13.97446	1.29704058	3
2	SISWA 2	0.890782	7.787617	5.049092978	1
3	SISWA 3	1.322897	9.534374	3.175656292	1
4	SISWA 4	3.115481	11.37574	1.455055347	3
5	SISWA 5	2.442212	10.76762	1.93616232	3
6	SISWA 6	3.517597	11.83314	0.955956447	3
7	SISWA 7	1.322897	9.534374	3.175656292	1
8	SISWA 8	2.778626	11.01654	1.79400895	3
9	SISWA 9	10.65235	2.2602	14.92881135	2
10	SISWA 10	7.849092	0.668568	12.11750192	2
11	SISWA 11	2.122691	10.52085	2.195860068	1
12	SISWA 12	1.029444	7.50915	5.183603971	1
13	SISWA 13	8.458906	1.303267	12.65668661	2
14	SISWA 14	3.339261	5.1067	7.648714079	1
15	SISWA 15	3.914046	12.31514	0.66865581	3
16	SISWA 16	8.330324	0.533576	12.58930178	2
17	SISWA 17	9.057781	0.656921	13.34347075	2
18	SISWA 18	10.65235	2.2602	14.92881135	2
19	SISWA 19	5.577096	13.95863	1.270442315	3
20	SISWA 20	6.235509	2.494009	10.54230418	2
21	SISWA 21	1.275946	7.522175	5.216247439	1
22	SISWA 22	6.055681	14.4391	1.750899083	3
23	SISWA 23	1.422175	7.53796	5.456800639	1
24	SISWA 24	7.962016	0.572693	12.26242637	2
25	SISWA 25	4.300406	5.044292	8.464284326	1
26	SISWA 26	8.811755	0.450109	13.09003365	2
27	SISWA 27	10.65235	2.2602	14.92881135	2
28	SISWA 28	5.605116	13.97446	1.29704058	3
29	SISWA 29	7.849092	0.668568	12.11750192	2
30	SISWA 30	1.322897	9.534374	3.175656292	1
31	SISWA 31	1.422175	7.53796	5.456800639	1
32	SISWA 32	1.422175	7.53796	5.456800639	1
33	SISWA 33	6.055681	14.4391	1.750899083	3
34	SISWA 34	1.275946	7.522175	5.216247439	1

Lanjutan tabel 4.7

No	NAMA	Jarak ke centroid 1	Jarak ke centroid 2	Jarak ke centroid 3	cluster
35	SISWA 35	1.322897	9.534374	3.175656292	1
36	SISWA 36	1.322897	9.534374	3.175656292	1
37	SISWA 37	7.849092	0.668568	12.11750192	2
38	SISWA 38	10.65235	2.2602	14.92881135	2
39	SISWA 39	1.322897	9.534374	3.175656292	1
40	SISWA 40	8.811755	0.450109	13.09003365	2
41	SISWA 41	3.914046	12.31514	0.66865581	3
42	SISWA 42	6.235509	2.494009	10.54230418	2
43	SISWA 43	8.811755	0.450109	13.09003365	2
44	SISWA 44	6.235509	2.494009	10.54230418	2
45	SISWA 45	8.811755	0.450109	13.09003365	2
46	SISWA 46	6.235509	2.494009	10.54230418	2
47	SISWA 47	1.322897	9.534374	3.175656292	1
48	SISWA 48	1.422175	7.53796	5.456800639	1
49	SISWA 49	2.122691	10.52085	2.195860068	1
50	SISWA 50	1.422175	7.53796	5.456800639	1
51	SISWA 51	2.122691	10.52085	2.195860068	1
52	SISWA 52	1.322897	9.534374	3.175656292	1
53	SISWA 53	3.914046	12.31514	0.66865581	3
54	SISWA 54	3.914046	12.31514	0.66865581	3

Dari hasil percobaan diatas dihasilkan kelompok siswa dengan pemahaman tinggi sebanyak 13 siswa, pemahaman sedang sebanyak 19 siswa, dan pemahaman kurang sebanyak 22 siswa.

4.1.6 Uji Coba Ke-6

Pada pengujian 1 sampai 5 data yang di gunakan adalah nilai rata-rata dari nilai pengetahuan dan nilai keterampilan. Pengujian ke-6 menggunakan data nilai dari mata pelajaran tanpa di rata-rata. Nilai yang digunakan adalah 12 nilai yaitu Quran Hadits pengetahuan, Quran Hadits keterampilan, Akidah Akhlak

pengetahuan, Akidah Akhlak keterampilan, Fikih pengetahuan, Fikih keterampilan, SKI pengetahuan, SKI keterampilan, Bahasa Arab pengetahuan, Bahasa Arab keterampilan, Tematik pengetahuan, Tematik keterampilan. Centroid yang digunakan sama dengan centroid pada uji coba 1 sampai 5.

Tabel 4. 8 Hasil uji coba 6

Nama	Klaster				
	Menggunakan Centroid Uji Coba 1	Menggunakan Centroid Uji Coba 2	Menggunakan Centroid Uji Coba 3	Menggunakan Centroid Uji Coba 4	Menggunakan Centroid Uji Coba 5
SISWA 1	3	3	3	3	3
SISWA 2	2	2	2	2	1
SISWA 3	2	2	2	2	1
SISWA 4	3	3	3	3	3
SISWA 5	3	3	3	3	3
SISWA 6	3	3	3	3	3
SISWA 7	2	2	2	2	1
SISWA 8	3	3	3	3	3
SISWA 9	1	1	1	1	2
SISWA 10	1	1	1	1	2
SISWA 11	2	2	2	3	1
SISWA 12	2	2	2	2	1
SISWA 13	1	1	1	1	2
SISWA 14	2	2	2	2	1
SISWA 15	3	3	3	3	3
SISWA 16	1	1	1	1	2
SISWA 17	1	1	1	1	2
SISWA 18	1	1	1	1	2
SISWA 19	3	3	3	3	3
SISWA 20	1	1	1	1	2
SISWA 21	2	2	2	2	1
SISWA 22	3	3	3	3	3
SISWA 23	2	2	2	2	1
SISWA 24	1	1	1	1	2
SISWA 25	2	2	2	2	1

Lanjutan tabel 4.8

Nama	Klaster				
	Menggunakan Centroid Uji Coba 1	Menggunakan Centroid Uji Coba 2	Menggunakan Centroid Uji Coba 3	Menggunakan Centroid Uji Coba 4	Menggunakan Centroid Uji Coba 5
SISWA 26	1	1	1	1	2
SISWA 27	1	1	1	1	2
SISWA 28	3	3	3	3	3
SISWA 29	1	1	1	1	2
SISWA 30	2	2	2	2	1
SISWA 31	2	2	2	2	1
SISWA 32	2	2	2	2	1
SISWA 33	3	3	3	3	3
SISWA 34	2	2	2	2	1
SISWA 35	2	2	2	2	1
SISWA 36	2	2	2	2	1
SISWA 37	1	1	1	1	2
SISWA 38	1	1	1	1	2
SISWA 39	2	2	2	2	1
SISWA 40	1	1	1	1	2
SISWA 41	3	3	3	3	3
SISWA 42	1	1	1	1	2
SISWA 43	1	1	1	1	2
SISWA 44	1	1	1	1	2
SISWA 45	1	1	1	1	2
SISWA 46	1	1	1	1	2
SISWA 47	2	2	2	2	1
SISWA 48	2	2	2	2	1
SISWA 49	2	2	2	3	1
SISWA 50	2	2	2	2	1
SISWA 51	2	2	2	3	1
SISWA 52	2	2	2	2	1
SISWA 53	3	3	3	3	3
SISWA 54	3	3	3	3	3

Pengujian dilakukan sebanyak 5 kali uji coba dengan *centroid* awal yang berbeda-beda dan 1 kali uji coba menggunakan input nilai yang berbeda. Pengujian dalam penelitian dilakukan dengan mengikuti ketentuan pembobotan data *K-means* secara default dan menentukan sendiri *centroid* awal. Hasil uji coba menampilkan *silhouette coefficient* untuk menilai kualitas kluster. Berikut merupakan hasil *Silhouette coefficient* dari 5 kali uji coba:

Tabel 4. 9 Hasil *Silhouette coefficient*

Uji coba ke-	Hasil <i>Silhouette coefficient</i>
1	0.5863
2	0.5863
3	0.5863
4	0.5901
5	0.5806

Tabel 4. 10 Hasil seluruh uji coba

	Uji Coba 1	Uji Coba 2	Uji Coba 3	Uji Coba 4	Uji Coba 5
SISWA 1	klaster 3				
SISWA 2	klaster 1				
SISWA 3	klaster 2	klaster 2	klaster 2	klaster 2	klaster 1
SISWA 4	klaster 3				
SISWA 5	klaster 3				
SISWA 6	klaster 3				
SISWA 7	klaster 2	klaster 2	klaster 2	klaster 2	klaster 1
SISWA 8	klaster 3				
SISWA 9	klaster 1	klaster 1	klaster 1	klaster 1	klaster 2
SISWA 10	klaster 1	klaster 1	klaster 1	klaster 1	klaster 2
SISWA 11	klaster 2	klaster 2	klaster 2	klaster 3	klaster 1
SISWA 12	klaster 2	klaster 2	klaster 2	klaster 2	klaster 1
SISWA 13	klaster 1	klaster 1	klaster 1	klaster 1	klaster 2
SISWA 14	klaster 2	klaster 2	klaster 2	klaster 2	klaster 1
SISWA 15	klaster 3				
SISWA 16	klaster 1	klaster 1	klaster 1	klaster 1	klaster 2
SISWA 17	klaster 1	klaster 1	klaster 1	klaster 1	klaster 2

Lanjutan tabel 4.10

	Uji Coba 1	Uji Coba 2	Uji Coba 3	Uji Coba 4	Uji Coba 5
SISWA 18	klaster 1	klaster 1	klaster 1	klaster 1	klaster 2
SISWA 19	klaster 3				
SISWA 20	klaster 1	klaster 1	klaster 1	klaster 1	klaster 2
SISWA 21	klaster 2	klaster 2	klaster 2	klaster 2	klaster 1
SISWA 22	klaster 3				
SISWA 23	klaster 2	klaster 2	klaster 2	klaster 2	klaster 1
SISWA 24	klaster 1	klaster 1	klaster 1	klaster 1	klaster 2
SISWA 25	klaster 2	klaster 2	klaster 2	klaster 2	klaster 1
SISWA 26	klaster 1	klaster 1	klaster 1	klaster 1	klaster 2
SISWA 27	klaster 1	klaster 1	klaster 1	klaster 1	klaster 2
SISWA 28	klaster 3				
SISWA 29	klaster 1	klaster 1	klaster 1	klaster 1	klaster 2
SISWA 30	klaster 2	klaster 2	klaster 2	klaster 2	klaster 1
SISWA 31	klaster 2	klaster 2	klaster 2	klaster 2	klaster 1
SISWA 32	klaster 2	klaster 2	klaster 2	klaster 2	klaster 1
SISWA 33	klaster 3				
SISWA 34	klaster 2	klaster 2	klaster 2	klaster 2	klaster 1
SISWA 35	klaster 2	klaster 2	klaster 2	klaster 2	klaster 1
SISWA 36	klaster 2	klaster 2	klaster 2	klaster 2	klaster 1
SISWA 37	klaster 1	klaster 1	klaster 1	klaster 1	klaster 2
SISWA 38	klaster 1	klaster 1	klaster 1	klaster 1	klaster 2
SISWA 39	klaster 2	klaster 2	klaster 2	klaster 2	klaster 1
SISWA 40	klaster 1	klaster 1	klaster 1	klaster 1	klaster 2
SISWA 41	klaster 3				
SISWA 42	klaster 1	klaster 1	klaster 1	klaster 1	klaster 2
SISWA 43	klaster 1	klaster 1	klaster 1	klaster 1	klaster 2
SISWA 44	klaster 1	klaster 1	klaster 1	klaster 1	klaster 2
SISWA 45	klaster 1	klaster 1	klaster 1	klaster 1	klaster 2
SISWA 46	klaster 1	klaster 1	klaster 1	klaster 1	klaster 2
SISWA 47	klaster 2	klaster 2	klaster 2	klaster 2	klaster 1
SISWA 48	klaster 2	klaster 2	klaster 2	klaster 2	klaster 1
SISWA 49	klaster 2	klaster 2	klaster 2	klaster 3	klaster 1
SISWA 50	klaster 2	klaster 2	klaster 2	klaster 2	klaster 1
SISWA 51	klaster 2	klaster 2	klaster 2	klaster 3	klaster 1

Lanjutan tabel 4.11

	Uji Coba 1	Uji Coba 2	Uji Coba 3	Uji Coba 4	Uji Coba 5
SISWA 52	klaster 2	klaster 2	klaster 2	klaster 2	klaster 1
SISWA 53	klaster 3				
SISWA 54	klaster 3				

4.2 Pembahasan

Pada penelitian ini sebelum melakukan pengujian, hal yang harus dilakukan pertama kali adalah mengumpulkan data. Pada penelitian ini data yang digunakan adalah berupa data nilai rapor kurikulum 2013 siswa-siswi kelas 3 tahun pelajaran 2023/2024 di MI Miftahul Uquliyah Kabupaten Malang yang terdiri dari nilai pengetahuan dan nilai keterampilan yang telah di konversi menjadi angka desimal sesuai bobot pada masing-masing kriteria.

Dari uji coba 1 sampai 5 memiliki hasil yang berbeda. Masing-masing uji coba memiliki centroid awal yang di tentukan. Nilai centroid yang ditentukan berpengaruh pada perhitungan jarak sehingga hasil jarak ke centroid serta pengelompokan kalster berbeda pada setiap uji coba. Pada uji coba 6 hasilnya adalah sama dengan uji coba 1 sampai 5. Nilai *silhouette score* yang di dihasilkan juga sama. Hasil dari lima kali uji coba mendapatkan nilai *silhouette score* yang berbeda. *Silhouette score* yang baik adalah dimana nilai tersebut mendekati 1. Nilai *silhouette score* yang paling mendekati 1 adalah pada uji coba ke-4. Nilai *silhouette score* yang mendekati 1 berarti *cluster* terpisah dengan baik dan data dalam satu *cluster* lebih mirip satu sama lain daripada dengan data pada *cluster* lain.

Hasil pada proses penentuan klaster ini tidak jauh berbeda dari penilaian yang di berikan oleh wali kelas kelas 3 MI Miftahul Uquliyah. Dari 54 siswa terdapat 10

siswa yang hasilnya berbeda dengan penilaian langsung dari wali kelas. Sistem pembelajaran yang saat ini digunakan siswa sering diberikan materi yang sama dalam satu kelas. Penyesuaian biasanya dilakukan hanya pada siswa yang menonjol. Hasil dari cluster yang dihasilkan oleh metode *k-means* dapat mengidentifikasi kelompok siswa yang memerlukan pendekatan belajar khusus, misalnya siswa dengan pemahaman tinggi yang memerlukan tantangan tambahan, atau siswa yang memerlukan bantuan lebih lanjut. Ini mendukung penerapan dimana materi dan metode pengajaran dapat disesuaikan untuk memenuhi kebutuhan setiap kelompok.

Berdasarkan dari hasil uji coba dapat disimpulkan bahwa algoritma *k-means* dapat menentukan pengelompokkan siswa berdasarkan tingkat pemahamannya. Oleh karena itu, manfaat yang guru dapat lebih optimal melakukan proses mengajar dengan baik, sehingga dapat menyesuaikan metode pengajaran dan strategi pembelajaran sesuai dengan karakteristik siswa. Hal ini akan meningkatkan efektivitas proses pembelajaran dan membantu guru dalam mencapai tujuan pendidikan yang diharapkan. Selain itu, sekolah dapat menyalurkan potensi siswa melalui keikutsertaan dalam lomba sesuai dengan kelompok pemahamannya. Penelitian ini juga diharapkan bermanfaat untuk para siswa agar lebih giat dalam menuntut ilmu karena pentingnya menuntut ilmu di dalam Islam sebagaimana sabda Rasulullah SAW :

أَلَا إِنَّ الدُّنْيَا مَلْعُونَةٌ مَلْعُونٌ مَا فِيهَا إِلَّا ذَكَرُ اللَّهِ وَمَا وَالَاهُ وَعَالِمٌ أَوْ مُتَعَلِّمٌ

"Ketahuilah bahwa sesungguhnya dunia itu terlaknat dan terlaknat pula isinya kecuali berdzikir kepada Allah dan ketaatan kepada-Nya, orang berilmu, dan orang yang belajar." (HR. At-Tirmidzi no. 2322).

Hadits diatas menjelaskan bahwa terdapat tiga amalan dimana jika orang tersebut meninggal maka pahalanya akan terus bertambah yaitu shadaqah jariyah, ilmu yang dimanfaatkan orang, atau anak shaleh yang mendoakan orangtuanya. Jika ilmu yang didapatkan para siswa bisa di amalkan oleh mereka sehingga bermanfaat bagi orang lain.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan penelitian dan pembahasan yang dilakukan, hasil dari penelitian adalah pengelompokan peserta didik berdasarkan nilai pengetahuan dan nilai keterampilan menggunakan algoritma *K-Means*. Hasil dari perhitungan tersebut yaitu mendapatkan 3 kelompok *cluster*. *Cluster 1* yaitu siswa dengan pemahaman tinggi sebanyak 16 siswa. *Cluster 2* yaitu siswa dengan pemahaman cukup sebanyak 19 siswa. *Cluster 3* yaitu siswa dengan pemahaman kurang yaitu sebanyak 19 siswa. Pengujian terhadap algoritma k-means dilakukan sebanyak 5 kali uji coba *centroid* dan 1 kali uji coba dengan input 12 nilai.

5.2 Saran

Dalam penelitian ini peneliti sadar bahwa masih banyak kekurangan. Untuk lebih mengembangkan penelitian ini, sebaiknya

1. Peneliti berharap penelitian di masa depan akan memperluas penggunaan metode selain *K-Means*, sehingga memungkinkan mereka membandingkan keefektifitasannya dengan metode yang digunakan dalam penelitian ini.
2. Diharapkan dapat mengimplementasikan pada sebuah tampilan sistem dan juga menambahkan fitur lain.

DAFTAR PUSTAKA

- Ferawati Jafar, A. (2021). *Print) Al asma: Journal of Islamic Education ISSN* (Vol. 3, Nomor 2).
- Ganga, S., Meyyappan, T., Phil, M., Karaikudi, K., & Tamilnadu, T. (2014). International Journal of Computer Science and Mobile Computing Performance of Students Evaluation in Education Sector Using Clustering K-Means Algorithms. Dalam *International Journal of Computer Science and Mobile Computing* (Vol. 3, Nomor 7). www.ijcsmc.com
- Hew, K. F., Jia, C., Gonda, D. E., & Bai, S. (2020). Transitioning to the “new normal” of learning in unpredictable times: pedagogical practices and learning performance in fully online flipped classrooms. *International Journal of Educational Technology in Higher Education*, 17(1). <https://doi.org/10.1186/s41239-020-00234-x>
- Khairani, N. A., & Sutoyo, E. (2020). Application of K-Means Clustering Algorithm for Determination of Fire-Prone Areas Utilizing Hotspots in West Kalimantan Province. *International Journal of Advances in Data and Information Systems*, 1(1), 9–16. <https://doi.org/10.25008/ijadis.v1i1.13>
- Komariyah, S., Anwar, S., & Nurhakim, B. (2023). *Implementasi Data Mining FP-Growth Untuk Analisis Pola Pembelian Pada Transaksi Penjualan*. 1(2).
- Kuleto, V., Ilić, M., Dumangiu, M., Ranković, M., Martins, O. M. D., Păun, D., & Mihoreanu, L. (2021). Exploring opportunities and challenges of artificial intelligence and machine learning in higher education institutions. *Sustainability (Switzerland)*, 13(18). <https://doi.org/10.3390/su131810424>
- Lia Hananto, A., Assiroj, P., Priyatna, B., Nurhayati, Fauzi, A., Yuniar Rahman, A., & Shofiah Hilabi, S. (2021). Analysis of Drug Data Mining with Clustering Technique Using K-Means Algorithm. *Journal of Physics: Conference Series*, 1908(1). <https://doi.org/10.1088/1742-6596/1908/1/012024>
- Macintyre, P. D., Gregersen, T., & Mercer, S. (2019). Setting an Agenda for Positive Psychology in SLA: Theory, Practice, and Research. *Modern Language Journal*, 103(1), 262–274. <https://doi.org/10.1111/modl.12544>
- Maria, H., Machado, D., & Vaz, R. (t.t.). *Master's degree Program in Data Science and Advanced Analytics DESCRIPTIVE ANALYSIS OF ONLINE ROULETTE GAMBLERS Segmentation of different gamblers based on their behavior using data mining algorithms*.

- Pérez-Ortega, J., Nely Almanza-Ortega, N., Vega-Villalobos, A., Pazos-Rangel, R., Zavala-Díaz, C., & Martínez-Rebollar, A. (t.t.). *The K-Means Algorithm Evolution*. www.intechopen.com
- Sagala, N. T. M., & Gunawan, A. A. S. (2022). Discovering the Optimal Number of Crime Cluster Using Elbow, Silhouette, Gap Statistics, and NbClust Methods. *ComTech: Computer, Mathematics and Engineering Applications*, 13(1), 1–10. <https://doi.org/10.21512/comtech.v13i1.7270>
- Sofiani, F., & Retno Utari, D. (t.t.). *ANALISIS DATA LOWONGAN KERJA SEKRETARIS PADA MASA PANDEMI COVID-19 DENGAN METODE DATA MINING KLAUSTERISASI* (Vol. 20, Nomor 2).
- Syahrnita, R., & Zaman, S. (2023). Regresi Logistik Multinomial untuk Prediksi Kategori Kelulusan Mahasiswa. Dalam *Jurnal Informatika Sunan Kalijaga* (Vol. 8, Nomor 2). MEI.
- Szymkowiak, A., Melović, B., Dabić, M., Jeganathan, K., & Kundi, G. S. (2021). Information technology and Gen Z: The role of teachers, the internet, and technology in the education of young people. *Technology in Society*, 65. <https://doi.org/10.1016/j.techsoc.2021.101565>
- Yuan, C., & Yang, H. (2019). Research on K-Value Selection Method of K-Means Clustering Algorithm. *J*, 2(2), 226–235. <https://doi.org/10.3390/j2020016>