

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1. Kelompok Belajar**

Menurut N. K. Roestiyah (2001) dalam kegiatan belajar-mengajar di kelas adakalanya guru membentuk kelompok kecil. Dalam pembelajaran kelompok kecil, guru memberikan bantuan atau bimbingan kepada tiap anggota kelompok secara lebih intensif. Hal ini supaya:

1. Hubungan belajar mengajar antara guru dan siswa menjadi lebih dekat.
2. Siswa memperoleh bantuan, kesempatan, sesuai dengan kebutuhan, kemampuan, dan minat.
3. Siswa dilibatkan dalam penentuan tujuan belajar, cara belajar, kriteria keberhasilan.[1]

Teknik pembelajaran dengan membentuk kelompok belajar merupakan salah satu strategi belajar mengajar, dimana siswa di dalam kelas dipandang sebagai suatu kelompok atau dibagi menjadi beberapa kelompok. Penggunaan teknik pembelajaran dengan kelompok belajar untuk proses belajar mengajar mempunyai tujuan agar siswa mampu bekerja sama dengan teman-teman yang lain dalam mencapai tujuan bersama dan dapat membangkitkan kegairahan siswa dalam belajar. Menurut Suprijono (2009), kelompok adalah dua individu atau lebih yang berinteraksi dan saling bergantung, yang saling bergabung untuk mencapai tujuan tertentu. Kelompok bukanlah semata-mata kumpulan orang yang saling berdekatan. Kelompok adalah kesatuan yang bulat diantara anggotanya. Sedangkan pembelajaran kelompok menurut Ibrahim (2000), merupakan pembelajaran yang dicirikan oleh struktur tugas, tujuan, dan penghargaan kelompok. Siswa bekerja dalam situasi pembelajaran kelompok didorong atau dikehendaki untuk bekerjasama pada suatu tugas dan mereka harus mengkoordinasi usahanya menyelesaikan tugasnya.

Tujuan dari kelompok belajar, yaitu:

1. Meningkatkan rasa percaya diri terhadap kemampuan siswa
2. Mengembangkan kemampuan siswa dalam bersosialisasi
3. Mewujudkan tingkah laku yang lebih efektif

4. Meningkatkan kemampuan berkomunikasi
5. Meningkatkan prestasi belajar siswa.

Adapun dasar pembagian kelompok belajar antara lain yaitu :

1. Kelompok belajar berdasarkan jumlah siswa per kelompok
  - a. Kelompok besar, dengan jumlah siswa antara 20-40 orang.
  - b. Kelompok kecil, dengan jumlah siswa antara 5-10 orang.
  - c. Kelompok individual, dengan jumlah siswa antara 1-5 orang
2. Kelompok belajar berdasarkan kemampuan siswa
  - a. Kelompok belajar siswa dengan kemampuan sedang atau dibawah rata-rata, adalah kelompok belajar yang dibentuk berdasarkan pada kemampuan siswa yang masih membutuhkan bimbingan dan dorongan secara utuh supaya kelompok tersebut berhasil.
  - b. Kelompok belajar dengan kemampuan cukup, adalah kelompok belajar yang dibentuk berdasarkan pada kemampuan siswa yang masih membutuhkan motivasi dan perhatian supaya berhasil untuk mencapai tujuan
  - c. Kelompok belajar dengan kemampuan baik atau diatas rata-rata, adalah kelompok belajar yang dibentuk berdasarkan pada kemampuan siswa yang sudah mulai mandiri dalam menyelesaikan tugasnya.

Dalam pembelajaran dengan menggunakan kelompok belajar siswa memiliki peran yang efektif untuk mendukung keberhasilan proses belajar mengajar. Metode yang tepat dalam menentukan atau membuat kelompok belajar siswa akan memberikan proses belajar yang menyenangkan bagi siswa dan diharapkan mampu meningkatkan hasil belajar siswa sehingga mendukung keberhasilan proses belajar mengajar.

## **2.2. Kemampuan Belajar Siswa**

Kemampuan berasal dari kata mampu yang berarti kuasa (bisa, sanggup) melakukan sesuatu, sedangkan kemampuan berarti kesanggupan, kecakapan, kekuatan (Kamus Besar Bahasa Indonesia, versi online/daring (dalam jaringan).

Kemampuan (*ability*) berarti kapasitas seorang individu untuk melakukan beragam tugas dalam suatu pekerjaan. (Stephen P. Robbins & Timonthy A. Judge, 2009: 57).

Dari pengertian tersebut dapat disimpulkan bahwa kemampuan adalah kesanggupan atau kecakapan seorang individu dalam menguasai suatu keahlian dan digunakan untuk mengerjakan beragam tugas dalam suatu pekerjaan.

Kemampuan juga dapat diartikan sebagai kompetensi sebagaimana dikemukakan oleh Stephen Robbin kompetensi adalah kemampuan (*ability*) kapasitas seseorang untuk mengerjakan berbagai tugas dalam suatu pekerjaan, dimana kemampuan ini ditentukan oleh dua faktor yaitu kemampuan kognitif dan kemampuan fisik. Menurut Van Looy, Van Dier dan Gemmel kompetensi adalah sebuah karakteristik manusia yang berhubungan dengan efektifitas performa, karakteristik ini dapat dilihat seperti gaya bertindak, berperilaku, dan berpikir. Dalam Undang-Undang No.13 Tahun 2013 kompetensi adalah kemampuan kerja setiap individu yang mencakup aspek pengetahuan, keterampilan dan sikap kerja yang sesuai dengan standar yang ditetapkan. (Muhammad Ridha Albaarr, 2020)

Dari beberapa pengertian para ahli dan Undang-Undang No.13 diatas maka dapat ditarik kesimpulan bahwa kompetensi adalah kemampuan, kecakapan dan karakteristik seseorang dalam menyelesaikan suatu pekerjaan berdasarkan intelektual, perilaku, cara berpikir, keterampilan serta sikap dalam pekerjaan untuk menyelesaikan suatu pekerjaan yang sesuai dengan standar yang telah ditetapkan.

Kompetensi dalam Kurikulum 13 muncul istilah Komptensi Inti sebagai unsur baru mengacu pada tujuan pendidikan. Dalam UU Sistem Pendidikan Nasional, Tujuan Pendidikan adalah: Berkembangnya potensi peserta didik agar menjadi manusia yang beriman dan bertakwa kepada Tuhan Yang Maha Esa, berakhlak mulia, sehat, berilmu, cakap, kreatif, mandiri, dan menjadi warga negara yang demokratis serta bertanggungjawab. Tujuan ini kemudian dikelompokkan ke dalam empat kelompok komptensi inti. [3] Empat kelompok kompetensi inti dapat dilihat pada tabel 3.1.

Tabel 2.1 Empat Kompetensi Inti

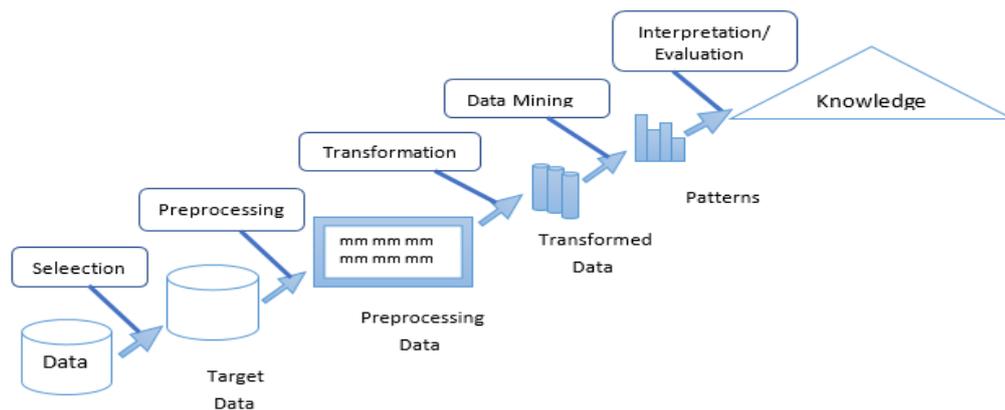
Tujuan Pendidikan	Dimensi	Kompetensi (KI)
Beriman dan bertakwa kepada Tuhan Yang Maha Esa	Sikap Spiritual	KI 1
Berakhlak mulia, sehat, mandiri, demokratis, bertanggungjawab	Sikap Sosial	KI 2
Berilmu	Pengetahuan	KI 3
Cakap dan kreatif	Keterampilan	KI 4

Kompetensi inti 1 dan 2 merupakan kompetensi generik yang harus dikuasai semua peserta didik tanpa kecuali. Hal ini sesuai dengan filosofi pendidikan dan dasar negara yaitu Pancasila. Apabila pendidikan tidak menghasilkan manusia yang memiliki ketakwaan kepada Tuhan dan tidak memiliki sikap sosial yang baik kepada sesama, maka dapat dikatakan bahwa pendidikan telah gagal.

Sedangkan dalam pencapaian Kompetensi Inti 3 dan 4 penguasaan peserta didik akan variatif sekali. Ilmu pengetahuan yang dikuasai peserta didik ada yang tinggi dan rendah. Demikian pula keterampilan yang dilakukan sesuai KI-4, setiap peserta didik akan memiliki kemampuan berbeda-beda.

### 2.3. *Knowledge Discovery in Database (KDD)*

*Knowledge Discovery in Database (KDD)* adalah proses menemukan pengetahuan yang berguna dari sebuah data yang bervolume besar, dan sering disebut sebagai *data mining*. KDD adalah proses yang terorganisir untuk mengidentifikasi pola-pola yang berlaku, berguna dan mudah dipahami dari kumpulan data yang besar dan kompleks. *Data mining* adalah inti dari proses KDD, yang melibatkan dalam menyimpulkan algoritma yang menjelajahi data, mengembangkan model dan menemukan pola-pola yang sebelumnya tidak diketahui. Model ini digunakan untuk memahami fenomena dari data, analisis dan prediksi[4]. Proses dalam *Knowledge Discovery in Database* dapat dilihat pada gambar 2.1 berikut ini:



Gambar 2.1 Proses Dalam KDD

Berikut penjelasan mengenai proses tahapan dalam *Knowledge Discovery in Database* (KDD), Ni Luh Wiwik, dkk, 2021):

### 1. Data Selection

Seleksi atau pemilihan data perlu dilakukan dari himpunan data operasional sebelum tahapan penggalian informasi dalam *Knowledge Discovery in Database* dimulai. Hasil data yang sudah diseleksi disimpan dalam suatu berkas dan akan dipergunakan dalam proses data mining, dimana letaknya terpisah dari basis data operasional (Kusrini & Lutfi, 2009)

### 2. Preprocessing

Tahapan pembersihan data ini merupakan operasi dasar sebelum proses data mining dilaksanakan. Proses dari cleaning ini meliputi beberapa aktivitas, seperti membuang duplikasi data, memeriksa data yang dinilai tidak konsisten serta mencari dan memperbaiki kesalahan yang ada pada data seperti kesalahan cetak (typography). Selain itu juga perlu dilaksanakan enrichment atau proses ‘memperkaya’ data sebelumnya dengan data yang lebih relevan serta diperlukan dalam KDD, contohnya data atau informasi yang bersifat eksternal (Kusrini & Lutfi, 2009)

### 3. Transformation

Menurut Fayyad (2009), tahapan transformation yaitu proses pencarian berbagai fitur yang dapat digunakan untuk mempresentasikan data, tergantung apa sasaran (goal) yang ingin dituju. Tahapan coding ini juga merupakan proses transformasi terhadap data yang diseleksi, hingga data tersebut dapat sesuai digunakan dalam proses data mining, dengan cara memodifikasi data ke dalam model analitis data

mining dan membuat model sesuai analisa yang diharapkan pada pola dan jenis informas yang dicar dalam database (basis data).

#### 4. Data Mining

Data mining di sini dapat diartikan sebagai proses menemukan pola dari informasi yang unik di dalam data yang telah diseleksi sebelumnya dengan menggunakan metode atau teknik tertentu. Adapun metode atau algoritma yang digunakan dalam data mining sendiri sangata variatif, di mana pemilihan teknik atau algoritma yang tepat sangat bergantung pada keseluruhan proses serta tujuan dari KDD (Kusrini & Lutfi, 2009)

#### 5. *Interpetation / Evaluation*

Tahapan ini adalah suatu bagian dari proses *Knowledge Discovery in Database* meliputi pemeriksaan terhadap pola ataupun informasi yang ditemukan, apakah bertentangan dengan hipotesa atau fakta yang telah ada sebelumnya. Pada tahapan *interpretation / evaluation* ini, juga dilakukan proses terjemahan dari beberapa pola yang didapatkan dari data mining. Pola informasi yang dihasilkan dari proses harus ditampilkan ke dalam bentuk yang lebih mudah dipahami oleh berbagai pihak yang berkepentingan (Fayyed, 1996).

#### 6. *Knowledge*

Tahap terakhir ini merupakan presentasi dari pola yang telah ditemukan untuk mendapatkan aksi pada tahap terakhir dari proses data mining ini adalah bagaimana cara memformulasikan aksi atau keputusan dari hasil analisis yang sudah didapatkan. Hasil mining yang telah melalui evaluasi digunakan dalam sistem informasi perusahaan dan perilaku organisasi.

### 2.4. Data Mining

Secara sederhana, *data mining* dapat diartikan sebagai proses mengekstrak atau menggali *knowledge* yang ada pada sekumpulan data. Informasi dan *knowledge* yang didapat tersebut dapat digunakan pada banyak bidang, seperti manajemen bisnis, pendidikan, kesehatan dan sebagainya. Berikut ini beberapa pengertian terkait data mining:

Menurut Cannolly dan Begg, 2010: “Data Mining adalah proses ekstraksi atau penggalian data yang belum diketahui sebelumnya, namun dapat dipahami dan

berguna dari database yang besar serta digunakan untuk membuat suatu keputusan bisnis yang sangat penting”. (Amril Mutoi Siregar, dkk, 2017)

Menurut Turban, dkk “Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Data mining merupakan proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengesktraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar”. (Kusrini dan Emha Taufiq Luthfi, 2009).

Menurut Pramudiono “Data mining merupakan serangkaian proses untuk menggali dan menadapatkan nilai tambah dari sekumpulan data yaitu berupa pengetahuan yang selama ini belum diketahui.” (Kusrini dan Emha Taufiq Luthfi, 2009).

Dari beberapa pengertian data mining diatas maka dapat disimpulkan bahwa data mining merupakan proses mencari pola atau informasi dari data terpilih menggunakan teknik atau metode tertentu untuk mendapatkan pengetahuan yang tersembunyi dari kumpulan data yang jumlahnya besar.

Data mining bertujuan untuk menemukan pola yang sebelumnya tidak diketahui. Jika pola-pola tersebut telah diperoleh maka dapat digunakan untuk menyelesaikan berbagai macam permasalahan. Data mining saat ini telah digunkan untuk membantu perusahaan fokus pada informasi penting dalam data yang telah dikumpulkan tentang perilaku pelanggan dan pelanggan potensial. Melalui data mining, perusahaan dapat menemukan informasi dalam data yang begitu besar melalui pengolahan yang tepat dan efektif.

Data mining secara umum selalu dikaitkan dengan bagaimana cara menyelesaikan permasalahan dengan menganalisis data yang terdapat dalam database. Data mining juga bisa dikualifikasikan sebagai suatu sains dan teknologi karena mampu memberikan sesuatu yang baru bagi penyelesaian masalah dengan teknik-teknik penyelesaian yang dipunyai.

## **2.5. Teknik-Teknik *Data mining***

*Data mining* adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara

manual. Perlu diingat bahwa kata *mining* sendiri berarti usaha untuk mendapatkan sedikit data berharga dari sejumlah besar data dasar. Karena itu *data mining* sebenarnya memiliki akar yang panjang dari bidang ilmu seperti kecerdasan buatan (*artificial intelligent*), *machine learning*, statistik dan basis data.

Menurut Ahmed, teknik *data mining* biasanya terbagi dalam dua kategori, prediksi dan deskripsi. Teknik prediksi menggunakan data historis untuk menyimpulkan sesuatu tentang kejadian di masa depan. Sedangkan teknik deskripsi bertujuan untuk menemukan pola dalam data yang menyediakan beberapa informasi tentang hubungan interval yang tersembunyi.

Menurut Kumar dan Saurabh, terdapat beberapa teknik yang digunakan dalam *data mining*, yaitu[4]:

### 1. *Classification*

Klasifikasi adalah teknik yang paling umum diterapkan pada *data mining*. Pendekatan ini sering menggunakan keputusan pohon (*decision tree*) atau *neural network* berbasis algoritma klasifikasi. Proses klasifikasi data melibatkan *learning* dan klasifikasi. Dalam belajar (*learning*) data pelatihan (*training*) dianalisis dengan algoritma klasifikasi. Dalam klasifikasi pengujian data dilakukan dengan menggunakan perkiraan akurasi dari aturan klasifikasi. Jika akurasi bisa diterima, maka aturan dapat diterapkan untuk data baru. Salah satu contoh yang mudah dan populer adalah dengan *decision tree* yaitu salah satu metode klasifikasi yang paling populer karena mudah untuk diinterpretasi. *Decision tree* adalah model prediksi menggunakan struktur pohon atau struktur berhirarki. *Decision tree* adalah struktur *flowchart* yang menyerupai *tree* (pohon), dimana setiap simpul internal menandakan suatu tes pada atribut, setiap cabang merepresentasikan hasil tes, dan simpul daun merepresentasikan kelas atau distribusi kelas. Alur pada *decision tree* di telusuri dari simpul akar ke simpul daun yang memegang prediksi kelas untuk contoh tersebut. *Decision tree* mudah untuk dikonversi ke aturan klasifikasi (*classification rules*)

### 2. *Clustering*

*Clustering* bisa dikatakan sebagai identifikasi kelas objek yang memiliki kemiripan. Dengan menggunakan teknik *Clustering* kita bisa lebih lanjut mengidentifikasi kepadatan dan jarak daerah dalam objek ruang dan dapat

menemukan secara keseluruhan pola distribusi dan korelasi antara atribut. Pendekatan klasifikasi secara efektif juga dapat digunakan untuk membedakan kelompok atau kelas objek

### 3. *Prediction*

Teknik regresi dapat disesuaikan untuk prediksi. Analisis regresi dapat digunakan untuk model hubungan antara satu atau lebih *independent variables* dan *dependent variables*. Dalam *data mining independent variabel* adalah atribut-atribut yang sudah dikenal dan respon variabel apa yang kita inginkan untuk diprediksi. Akan tetapi, banyak masalah di dunia nyata bukan prediksi yang mudah. Karena itu, teknik kompleks (seperti: *logistic regression*, *decision trees* atau pohon keputusan, *neural nets* atau jaringan syaraf) mungkin akan diperlukan untuk memprediksi nilai. Model yang berjenis sama sering dapat digunakan untuk regresi dan klasifikasi. Misalnya, CART (*Classification and Regression Trees*) yaitu algoritma pohon keputusan yang dapat digunakan untuk membangun ke-2 pohon klasifikasi dan pohon regresi. Jaringan saraf juga dapat menciptakan ke-2 model klasifikasi dan regresi

### 4. *Association rule*

Digunakan untuk mengenali kelakuan dari kejadian-kejadian khusus atau proses dimana link asosiasi muncul pada setiap kejadian. Contoh dari aturan assosiatif dari analisa pembelian di suatu pasar swalayan adalah bisa diketahui berapa besar kemungkinan seorang pelanggan membeli roti bersamaan dengan susu. Dengan pengetahuan tersebut pemilik pasar swalayan dapat mengatur penempatan barangnya atau merancang kampanye pemasaran dengan memakai kupon diskon untuk kombinasi barang tertentu.

Penting tidaknya suatu aturan assosiatif dapat diketahui dengan dua parameter, *support* yaitu prosentasi kombinasi atribut tersebut dalam basisdata dan *confidence* yaitu kuatnya hubungan antar atribut dalam aturan asosiatif. Motivasi awal pencarian *association rule* berasal dari keinginan untuk menganalisa data transaksi supermarket, ditinjau dari perilaku *customer* dalam membeli produk. *Association rule* ini menjelaskan seberapa sering suatu produk dibeli secara bersamaan.

### 5. *Neural network*

Jaringan saraf adalah seperangkat unit penghubung *input* dan *output* dimana setiap koneksinya memiliki bobot. Selama fase *learning*, jaringan belajar dengan menyesuaikan bobot sehingga dapat memprediksi kelas yang benar label dari setiap *input*. Jaringan saraf memiliki kemampuan yang luar biasa untuk memperoleh arti dari data yang rumit atau tidak tepat dan dapat digunakan untuk mengambil pola-pola serta mendeteksi *tren* yang sangat kompleks untuk diperhatikan baik oleh manusia atau teknik komputer lain. Jaringan saraf sangat baik untuk mengidentifikasi pola atau *tren* pada data dan sangat cocok untuk melakukan prediksi serta memprediksi kebutuhan.

### 6. *Decision trees*

*Decision trees* atau pohon keputusan adalah struktur *tree-shaped* yang mewakili set keputusan. Keputusan ini menghasilkan aturan untuk klasifikasi sebuah kumpulan data. Metode pohon keputusan diantaranya yaitu *Classification and regression trees* (CART) dan *Chi Square Automatic Interaction Detection* (CHAID).

### 7. *Nearest Neighbor Method*

Teknik yang mengklasifikasikan setiap *record* dalam sebuah kumpulan data berdasarkan sebuah kombinasi suatu kelas  $k$  *record* yang sama dalam sebuah kumpulan data historis (dimana  $k$  lebih besar atau sama dengan 1 ). Terkadang disebut juga dengan teknik *K-Nearest Neighbor*.

## 2.6. Metode *Clustering*

Madhu Yedha mendefinisikan *Clustering* sebagai proses pengorganisasian objek data ke dalam set kelas yang saling berhubungan, yang disebut *Cluster*. *Clustering* merupakan contoh dari klasifikasi tanpa arahan (unsupervised). Unsupervised berarti bahwa pengelompokan tidak tergantung pada standar kelas dan pelatihan atau training. [4]

Menurut Deka, *Clustering* merupakan salah satu teknik data mining yang digunakan untuk mendapatkan kelompok dari objek-objek yang mempunyai karakteristik yang umum di data yang cukup besar. Tujuan utama dari metode *Clustering* adalah pengelompokan sejumlah data atau objek ke dalam *Cluster* atau

grup sehingga dalam setiap *Cluster* akan berisi data yang semirip mungkin. *Clustering* melakukan pengelompokan data yang didasarkan pada kesamaan antar objek, oleh karena itu *cluterisasi* digolongkan sebagai metode unsupervised learning. Menurut Oyelade, *Clustering* dapat dibagi menjadi dua, yaitu hierarchical *Clustering* dan non-hierarchical *Clustering*. [4]

Hierarchical *Clustering* adalah suatu metode pengelompokan data yang dimulai dengan mengelompokkan dua atau lebih objek yang memiliki kesamaan paling dekat. Kemudian proses diteruskan ke objek lain yang memiliki kedekatan ke-2. Demikian seterusnya sehingga *Cluster* akan membentuk semacam pohon dimana ada hierarki (tingkatan ) yang jelas antar objek, dari yang paling mirip sampai yang paling tidak mirip. Secara logika semua objek pada akhirnya hanya akan membentuk sebuah *Cluster*. Dendogram biasanya digunakan untuk membantu memperjelas proses hierarki tersebut. [4]

Berbeda dengan metode hierarchical *Clustering*, metode non-hierarchical *Clustering* justru dimulai dengan menentukan terlebih dahulu jumlah *Cluster* yang diinginkan (dua *Cluster*, tiga *Cluster*, atau lain sebagainya). Setelah jumlah *Cluster* diketahui, baru proses *Cluster* dilakukan tanpa mengikuti proses hierarki. Metode ini biasa disebut dengan *K-Means Clustering*. [4]

## 2.7. Algoritma *K-Means Clustering*

Menurut Ika Kurniawati, dkk (2017:72) didalam jurnalnya: “*K-Means Clustering* merupakan salah satu metode data *Clustering* non-hirarki yang mengelompokkan data dalam bentuk satu atau lebih *Cluster*/kelompok. Data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu *Cluster*/kelompok dan data yang memiliki karakteristik yang berbeda dikelompokkan dengan *Cluster*/kelompok yang lain sehingga data yang berada dalam satu *Cluster*/kelompok memiliki tingkat variasi yang kecil. (Agusta, 2007). [5]

Langkah-langkah melakukan *Clustering* dengan metode K-Means adalah sebagai berikut:

1. Tentukan jumlah *Cluster* (*k*) yang diinginkan pada dataset.
2. Tentukan titik pusat (*Centroid*) secara acak/random pada tahap awal

3. Alokasikan semua data/ objek ke *Cluster* terdekat. Kedekatan dua objek ditentukan berdasarkan jarak ke-2 objek tersebut. Demikian juga kedekatan suatu data ke *Cluster* tertentu ditentukan jarak antara data dengan pusat *Cluster*. Dalam tahap ini perlu dihitung jarak tiap data ke tiap pusat *Cluster*. Jarak paling dekat antara satu data dengan satu *Cluster* tertentu akan menentukan suatu data masuk dalam *Cluster* mana. Untuk menghitung jarak semua data ke setiap titik pusat *Cluster* dapat menggunakan teori jarak Euclidean dengan rumus sebagai berikut:

$$D(i, j) = \sqrt{(x_{1i} - x_{1j})^2 + (x_{2i} - x_{2j})^2 + \dots + (x_{ki} - x_{kj})^2} \dots (1)$$

Keterangan:

$D(i, j)$  = Jarak antara ke  $i$  ke pusat *Cluster*  $j$

$X_{ki}$  = Data ke  $i$  pada atribut data ke  $k$

$X_{kj}$  = Titik pusat ke  $j$  pada atribut ke  $k$

4. Hitung kembali pusat *Cluster* dengan keanggotaan *Cluster* yang sekarang. Pusat *Cluster* adalah rata-rata dari semua data/ objek dalam *Cluster* tertentu.
5. Hitung kembali setiap objek menggunakan pusat *Cluster* baru (*Centroid* baru), ini merupakan tahap awal pembukaan iterasi baru. Jika anggota *Cluster* tidak mengalami perpindahan *Cluster* lagi, maka proses *Clustering* dinyatakan selesai. Tetapi jika anggota *Cluster* mengalami perpindahan, maka kembali lagi pada langkah nomor 3 sampai anggota *Cluster* tidak mengalami perpindahan lagi”

## 2.8. RapidMiner

Rapidminer merupakan salah satu perangkat lunak yang dapat digunakan untuk mengolah data dengan menggunakan prinsip dan algoritma data mining. Rapidminer mengekstrak pola-pola dari dataset yang besar yang dikombinasikan metode statistika, kecerdasan buatan, dan database.

Dengan Rapidminer pengguna dimudahkan untuk melakukan perhitungan data yang jumlahnya banyak menggunakan operator-operator. fungsi operator-operator adalah untuk memodifikasi data. Data dihubungkan dengan node-node pada operator kemudian pengguna dapat melihat hasilnya. Hasil yang diperlihatkan dapat ditampilkan secara visual menggunakan grafik. Hal ini yang menjadikan

rapidminer sebagai software pilihan untuk melakukan ekstraksi data dengan metode-metode data mining (Novianti, 2019).[2]

## 2.9. Penelitian Terdahulu

Dalam melakukan sebuah penelitian diperlukan kajian dari penelitian sebelumnya. Jurnal dari penelitian sebelumnya dapat menjadi bahan acuan membantu memahami permasalahan dan juga pemecahannya, serta menjadi pembandingan dengan penelitian yang dilakukan. Beberapa penelitian terdahulu yang peneliti gunakan sebagai *literatur* dalam penelitian ini yaitu jurnal yang berkaitan dengan *Algoritma K-Means Clustering*, dapat dilihat pada tabel 2.2 adalah sebagai berikut:

Tabel 2.2 Penelitian Terdahulu

No	Judul	Peneliti (Tahun)	Metode	Hasil Penelitian
1	Klaterisasi Siswa Berdasarkan Hasil Belajar Menggunakan K-Means Berbasis Web.[6]	Wiwit Agus Lestari, Kurnia Paranita Kartika, Saiful Nur Budiman (2022)	K-Means Clustering	Algoritma Kmeans terbukti mampu melakukan <i>cluterisasi</i> siswa berdasarkan hasil belajar untuk pembentukan kelompok belajar dengan hasil pengujian aplikasi menggunakan Silhouette Coefficient tertinggi sebesar 0,758. Keberhasilan sistem juga dinilai dari Black Box Testing yang menghasilkan tingkat keberhasilan 100% untuk seluruh menu yang ada dalam aplikasi.

				Hasil pengujian pengguna terhadap aplikasi <i>clustering</i> juga sangat baik dengan nilai 97,84 %
2	Implementasi Algoritma <i>K-Means Clustering</i> Untuk Menentukan Kelas Kelompok Bimbingan Belajar Tambahan[7]	Mardalius (2017)	Algoritma <i>K-Means Clustering</i>	Metode <i>Clustering</i> dengan algoritma <i>K-Means</i> dapat digunakan untuk mengelompokkan data siswa berdasarkan nilai ujian, yaitu kemampuan siswa pintar, siswa sedang dan siswa kurang pintar sehingga dapat mengetahui siswa yang mana saja yang akan diberi belajar tambahan agar dapat mencapai nilai standar kelulusan Ujian Nasional dan membantu pihak sekolah untuk melakukan belajar tambahan terhadap siswa yang akan mengikuti Ujian Nasional
3	Penerapan Metode <i>Kmeans</i>	Venny Novita Sari,	<i>K-Means Clustering</i>	Dapat mengetahui kelompok mahasiswa yang memiliki nilai rata-rata IPK tertinggi dari

	<i>Clustering</i> dalam Menentukan predikat Kelulusan mahasiswa Untuk menganalisa kualitas lulusan[8]	Yupianti, Dewi Maharani (2018)		ketiga <i>Cluster</i> terdapat pada <i>Cluster 2</i> dengan rata – rata IPK 3.3967 sehingga dapat disimpulkan bahwa <i>Cluster 2</i> merupakan lulusan mahasiswa yang memiliki kualitas terbaik
4	<i>Cluterisasi</i> Penempatan Siswa yang Optimal untuk Meningkatkan Nilai Rata-Rata Kelas Menggunakan K-Means. [9]	Elda, Yusma Defit, Sarjon Yunus, Yuhandri Syaljumairi, Raemon (2021)	K-Means Clustering	Hasil <i>cluterisasi</i> data 90 siswa dikelompokkan menjadi 3 <i>Cluster</i> dengan jumlah data siswa untuk <i>Cluster 1</i> dengan bobot sikap dan hasil belajar tinggi berjumlah 47 siswa, <i>Cluster 2</i> dengan bobot sikap dan hasil belajar sedang berjumlah 10 siswa dan <i>Cluster 3</i> dengan bobot sikap dan hasil belajar rendah berjumlah 33 siswa.
5	Sistem Informasi Pemetaan Wilayah	Nurjoko, Defi Dwirohayati, Novi	<i>K-Means Clustering</i>	SIG yang diperoleh dari penggabungan semua data spasial berdasarkan hasil <i>Clustering</i> dapat

	Rawan Kriminalitas Polresta Bandar Lampung Menggunakan <i>K-Means Clustering</i> [10]	Herawadi Sudibyo (2020)		merekomendasikan pihak berwenang untuk menindaklanjuti wilayah yang memiliki intensitas kejahatan tinggi, membantu dalam pembuatan laporan pihak berwenang dan website akan menginformasikan kepada masyarakat di Kota Bandar Lampung, dimana lokasi daerah rawan kejahatan dan lokasinya aman
6	Implementasi K-Means Clushtering Dalam Mengelompokk an Minat Membaca Penduduk Menurut Wilayah[11]	Mhd Yuda Rizki, Siti Maysaroh, dan Agus Perdana Windarto (2021)	K-Means Clushtering	Dengan menggunakan algoritma Kmeans clushtering dapat menjadi salah satu pilihan terbaik dalam kasus mengelompokkan. Hasil dari penelitian ini menghasilkan -12 Provinsi clushter tingkat tinggi(C1) dan -21 Provinsi clushter tingkat rendah (C2). Dengan memanfaatkan software RapidMiner dapat menjadi salah satu pilihan terbaik dalam mengji hasil penelitian

menggunakan algoritma  
K-Means clustering

---

Pada tabel 2.1. Ditampilkan hasil penelitian sebelumnya yang semuanya menggunakan metode *K-Means Clustering*. Dari beberapa penelitian yang telah dilakukan menyebutkan bahwa *Clustering* menggunakan algoritma K-Means relatif dapat digunakan untuk mengklaster data dengan hasil klaster yang berkualitas. Maka dalam penelitian ini peneliti menggunakan *K-Means Clustering* untuk menentukan kelompok belajar siswa berdasarkan kemampuan. Untuk hal ini Peneliti melakukan klastering kemampuan belajar siswa di lima mata pelajaran yang biasanya digunakan untuk Ujian Nasional, yaitu: Bhs. Indonesia, Matematika, IPA, IPS, dan Bhs. Inggris. Terdapat tiga klaster untuk kemampuan siswa dimasing-masing mata pelajaran tersebut antara lain kemampuan belajar sedang, cukup, dan baik. Untuk melakukan Analisis dan uji performa menggunakan Aplikasi Rapidminer