

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

Setiap sektor, termasuk perbankan, kesehatan, bisnis, dan bisnis, telah terpengaruh oleh perkembangan dan penerapan teknologi informasi yang semakin pesat. Selain itu, teknologi informasi telah memengaruhi pendidikan, terutama proses pembelajaran (Maghfiroh, 2020). Menurut Cintamulya (2012), pembelajaran yang diberikan harus mampu meningkatkan kemampuan sumber daya manusia untuk menyiapkan profil lulusan yang memiliki kemampuan yang relevan dengan era teknologi atau zaman baru. Perguruan tinggi dan sekolah adalah salah satu sumber daya yang dapat mempersiapkan mahasiswa dan murid sebagai sumber daya manusia yang handal untuk menghadapi era modern berbasis teknologi dan globalisasi.

Dalam peran mereka sebagai penyelenggara pendidikan, sekolah sering menghadapi tantangan karena tidak mencapai visi dan misi mereka. Ada banyak alasan mengapa hasil yang diinginkan tidak tercapai. Faktor internal termasuk sumber daya manusia, metode pengajaran, dan bahkan kurikulum yang tidak memenuhi standar (Setianingrum, 2014). Permasalahan lainnya, siswa yang baru memasuki sekolah rentan terhadap banyaknya tantangan yang mempengaruhi kinerja mereka dikarenakan menghadapi lingkungan yang baru (Saeed & Mohamedali, 2022). Evaluasi kinerja siswa memiliki peranan yang sangat penting di lembaga-lembaga pendidikan, khususnya di sekolah. Metode monitoring dan evaluasi yang dapat bersifat subjektif dan obyektif menjadi indikator tolak ukur dalam proses pembelajaran dengan target dengan harapan dapat memberikan saran untuk bimbingan konseling pada siswa yang bertujuan untuk dapat memperbaiki nilai (Ulinnuha, 2016).

Kinerja siswa menjadi faktor kunci terhadap kemampuan pengajar serta memberikan reputasi yang baik atau buruk bagi sekolah (Saleh et al., 2021). Siswa setiap semester mendapatkan nilai yang menjadi tolak ukur perkembangannya.

Parameter khusus ini digunakan untuk menentukan kualitas dan untuk memantau siswa sejak dini untuk memperbaiki nilai. Sangat penting untuk melakukan penilaian siswa berdasarkan kelompokan karena bertujuan untuk memantau perkembangan siswa (Ulinnuha, 2016). Analisis kelompokan, juga dikenal sebagai analisis kelompokan, adalah analisis yang bermanfaat untuk mengelompokkan item berdasarkan kemiripan parameter-parameter yang diamati (Inyang & Joshua, 2013).

Permasalahan utama dalam melakukan evaluasi kinerja siswa adalah pengelompokan, di mana kelompok dibentuk berdasarkan tingkat kecerdasan dan jumlah siswa tidak boleh melebihi kapasitas yang telah ditentukan. Pengelompokan berdasarkan tingkat kecerdasan sangat penting untuk menjaga homogenitas kelompok. Kelompok yang tidak homogen akan sulit untuk memberikan layanan pendidikan yang baik kepada populasi siswa yang sangat beragam. Evaluasi kinerja siswa melibatkan beberapa komponen yang didasarkan pada jumlah penilaian yang terkadang tidak tepat karena interpretasi manusia. Selain itu, metode aritmatika dan statistik yang umum digunakan masih memiliki beberapa keterbatasan (Patel & Yadav, 2015).

Pada penelitian ini mencoba menyajikan studi tentang data mining pada bidang pendidikan untuk memodelkan kinerja siswa. Dalam bidang pendidikan, data mining adalah salah satu metode yang digunakan untuk melakukan analisis data yang diperoleh dari hasil proses pembelajaran. Tujuannya adalah untuk mendapatkan informasi baru dari data yang sudah ada dan dengan demikian meningkatkan kualitas pembelajaran. Pengelompokan kinerja siswa berdasarkan potensi adalah merupakan cara yang dilakukan untuk proses peningkatan pembelajaran yang berkualitas. Pengelompokan kinerja siswa berdasarkan potensi dapat menjadi strategi untuk meningkatkan prestasi belajar siswa, pengelolaan data siswa yang lebih baik, dan meningkatkan kelulusan mereka (Wulandari et al., 2020).

Salah satu penelitian yang dilakukan oleh (Ulinnuha, 2016) menggunakan metode pengelompokan *Fuzzy C-Means Clustering* untuk mengklasifikasikan kinerja siswa. Nanti, setiap parameter akan dibagi menjadi kelompok siswa dengan

nilai yang sangat memuaskan, memuaskan, dan kurang memuaskan. Tujuannya adalah untuk menjadi lebih mudah untuk melacak siswa yang berada dalam kelompok yang sama. Adanya sistem pemantauan akan membantu siswa memperbaiki hasil akademiknya. Hasil uji coba yang dilakukan dengan algoritma *Fuzzy C-Means Clustering* didapatkan akurasi *training* dengan nilai RMSE sebesar  $2.76 \times 10^{-15}$  dan didapatkan akurasi *testing* dengan nilai RMSE sebesar 0,1165.

Penelitian yang dilakukan oleh (N. Divanadia Luckyana et all, 2021) menghasilkan penggabungan *Self-Organizing Map* (SOM) dan *Fuzzy C-Means* (FCM) mampu memberikan hasil performa kinerja proses evaluasi dengan cara menempatkan siswa kedalam kelompok yang sesuai agar dapat membantu tenaga pengajar dalam menangani serta memberikan umpan balik yang tepat pada siswanya. Diperoleh hasil proses evaluasi dengan menggunakan pengujian Wilcoxon test menunjukkan bahwa nilai t stat lebih kecil daripada nilai critical value yang berarti terdapat perbedaan yang signifikan antara hasil pengelompokan dengan SOM dan SOM-FCM. Pada penelitian ini juga menggunakan sumber data dari log data media pembelajaran digital mengelompokkan siswa berdasarkan aktivitas mereka selama belajar dengan media tersebut.

Penelitian lainnya yang dilakukan oleh (S. Kurniawan, et all, 2023) melakukan pengelompokan terhadap kinerja siswa prestasi berdasarkan nilai akademik berbasis konvensional dan komputerisasi pemrograman. Adapun algoritma yang digunakan yaitu K-Means Clustering dan Fuzzy C-Means. indikator pengelompokan ada 3 yaitu siswa berprestasi, prestasi sedang dan tidak berprestasi. Dari pengujian yang dihasilkan bahwa algoritma K-Means Clustering dan Fuzzy C-Means dengan metode davies bouldien index (DBI) menunjukkan bahwa setiap perhitungan merupakan yang terbaik karena memiliki nilai DBI yang mendekati 0 dan tidak negatif.

Prinsip kerja *Fuzzy C-Means Clustering* pada dasarnya mirip dengan teknik pengelompokan data lainnya. Data dengan karakteristik yang sama atau berdekatan akan dikategorikan pada *cluster* yang sama. Metode ini dapat digunakan dengan baik dan memiliki hasil yang lebih baik. Hasil pengelompokan kinerja menggunakan metode *Fuzzy C-Means Clustering* memiliki kinerja yang lebih baik

daripada pengelompokan yang dibentuk dengan metode, seperti: *K-Means Clustering*, *Fuzzy Logic*, dan Bayesian (Patel & Yadav, 2015). Namun, permasalahan umum dalam penggunaan teknik ini, seperti: sensitivitas tinggi terhadap kondisi awal (pengaturan awal pusat *cluster*), tidak ada jaminan untuk mencapai solusi optimal global, dan konvergensi lambat untuk memisahkan *cluster* (Xianfeng & Pengfei, 2015). Selain itu, masalah fitur yang digunakan dalam proses pengelompokan data sering diabaikan, sehingga fitur yang tidak penting digunakan dalam proses klusterisasi data. Terdapat beberapa penelitian yang telah melakukan kombinasi untuk perbaikan dan penyelesaian masalah pada metode *Fuzzy C-Means Clustering* (Hadi et al., 2016).

Penelitian yang dilakukan oleh (Hadi et al., 2016), mengkombinasikan metode *Fuzzy C-Means Clustering* dengan Algoritma Genetika dalam penentuan kinerja siswa. Hanya 61 dari 77 fitur dataset yang relevan dan dapat digunakan dalam proses clustering, dihasilkan dari kombinasi kedua algoritma ini: genetika dan fuzzy C-Means. Percobaan yang menggabungkan Algoritma Genetika dan Fuzzy C-Means menunjukkan hasil yang lebih baik daripada yang dihasilkan oleh Fuzzy C-Means saja, dengan persentase rata-rata kesesuaian sebesar 88,89%. Penelitian lainnya dilakukan oleh (Hossen et al., 2011), menggabungkan metode FCM dengan *subtractive clustering* (SC) untuk meningkatkan akurasi pengelompokan. Penelitian tersebut menunjukkan bahwa algoritma yang diusulkan mampu meningkatkan kecepatan, mengurangi iterasi, dan membuat partisi data lebih stabil serta akurasi lebih tinggi.

Berdasarkan penelitian yang telah dijelaskan sebelumnya, perbaikan atau peningkatan algoritma *Fuzzy C-Means Clustering* untuk pengelompokan data perlu dilakukan agar mendapatkan hasil yang lebih baik, terutama dalam penentuan pusat *cluster* awal yang tepat. Dengan menggunakan algoritma ROCK (RObust Clustering using linKs), penelitian ini akan mencoba meningkatkan atau memperbaiki Clustering Fuzzy C-Means. ROCK adalah algoritma clustering agglomeratif hierarchial untuk data kategori. Untuk data kategori, algoritma dengan pengelompokan hierarki tradisional menggunakan jarak antara titik dengan titik lainnya untuk pengelompokan dianggap sudah tidak efektif. Akibatnya,

diciptakan ide baru, yaitu link, yang difungsikan untuk mengukur kesamaan antara dua titik data (Dewi et al., 2021).

Pada penelitian ini, pengujian algoritma ROCK digunakan dalam penentuan pusat *cluster* awal pada algoritma *Fuzzy C-Means Clustering*. Algoritma ROCK akan mencari *cluster* data pada dataset dengan melakukan pengukuran jarak diantara data tersebut. Hasil dari *cluster* data tersebut akan dijadikan sebagai inisialisasi awal pada algoritma *Fuzzy C-Means Clustering* dalam melakukan pengelompokan data dan menentukan hasil kinerja siswa. algoritma ini dapat lebih fleksibel dalam menangani data yang kompleks atau tidak teratur, serta memberikan hasil clustering yang lebih informatif dengan mempertimbangkan tingkat keanggotaan fuzzy dari setiap titik data terhadap cluster-cluster yang ada. Ini membuatnya cocok untuk aplikasi di mana data memiliki struktur yang kompleks dan tidak dapat dijelaskan dengan baik oleh metode clustering yang lebih tradisional seperti k-means. Berdasarkan uraian sebelumnya, peneliti tertarik melakukan penelitian yang lebih mendalam terkait dengan peningkatan kinerja metode *Fuzzy C-Means Clustering* menggunakan algoritma ROCK, sehingga dapat meningkatkan hasil evaluasi kinerja siswa serta akurasi dalam pengelompokan data.

## 1.2. Rumusan Masalah

Pada penelitian ini sesuai dengan latar belakang yang telah disampaikan diatas, maka dapat dirumuskan masalah sebagai berikut, yaitu:

1. Bagaimana pemodelan pengelompokan kinerja siswa menggunakan algoritma *Fuzzy C-Means Clustering*.
2. Bagaimana pemodelan pengelompokan kinerja siswa menggunakan metode *Fuzzy C-Means Clustering* dan algoritma ROCK.
3. Bagaimana melakukan peningkatan kinerja atau perbaikan pada algoritma *Fuzzy C-Means Clustering* menggunakan algoritma ROCK untuk meningkatkan akurasi pengelompokan kinerja siswa.

## 1.3. Tujuan dan Manfaat Penelitian

### 1.3.1. Tujuan Penelitian

Pada penelitian ini sesuai dengan latar belakang yang telah diuraikan diatas, maka tujuan dari penelitian ini adalah sebagai berikut, yaitu:

1. Merancang pemodelan pengelompokan kinerja siswa menggunakan algoritma *Fuzzy C-Means Clustering*.
2. Merancang pemodelan pengelompokan kinerja siswa menggunakan metode *Fuzzy C-Means Clustering* dan algoritma ROCK.
3. Melakukan peningkatan kinerja atau perbaikan pada algoritma *Fuzzy C-Means Clustering* menggunakan algoritma ROCK untuk meningkatkan akurasi pengelompokan kinerja siswa.

### 1.3.2. Manfaat Penelitian

Adapun manfaat dari penelitian ini yang dilakukan adalah sebagai berikut, yaitu:

1. Mengetahui hasil kinerja siswa dan mempermudah sekolah melakukan monitoring dan evaluasi terhadap pencapaian hasil siswa.
2. Mengetahui hasil pengelompokan kinerja siswa pada suatu sekolah menggunakan metode *Fuzzy C-Means Clustering* dan algoritma ROCK.
3. Dengan adanya penelitian ini, dapat menjadi bahan referensi mengajar dan belajar, juga sebagai referensi untuk kasus atau pembahasan lainnya yang berkaitan dengan penelitian ini.

### 1.4. Pembatasan Masalah

Berdasarkan hasil pemaparan penelitian ini maka perlu diberikan batasan yang jelas dan dapat dicapai sehingga ruang lingkupnya hanya meliputi sebagai berikut, yaitu:

1. Pengambilan dan pengujian data diperoleh dari penilaian siswa di sekolah dan disesuaikan dengan fitur yang terdapat pada UCI *Machine Learning Repository*, yaitu *Student Performance Dataset*.
2. Pengelompokan kinerja siswa dilakukan menggunakan algoritma *Fuzzy C-Means Clustering* dan tidak membandingkannya dengan teknik *clustering* lain.

3. Perbaikan atau peningkatan hasil pengelompokan data pada algoritma *Fuzzy C-Means Clustering* akan dilakukan menggunakan algoritma ROCK dan tidak membandingkannya dengan teknik lain.
4. Pengujian hasil pengelompokan menggunakan algoritma *Fuzzy C-Means Clustering* dan algoritma ROCK akan dilihat dari peningkatan akurasi, *rand index*, dan *F-measure* yang diperoleh dan membandingkannya dengan akurasi, *rand index*, dan *F-measure* pengelompokan menggunakan algoritma *Fuzzy C-Means Clustering* biasa.

### 1.5. Sistematika Penulisan

Pembahasan pada penelitian ini dibagi menjadi 5 (lima) bab, dengan tujuan untuk pembahasan sesuai prosedur dan sistematis. Penjelasan secara singkat mengenai masing-masing bab dapat dijelaskan sebagai berikut, yaitu:

1. Bab I Pendahuluan  
Berisi tentang Latar Belakang, Rumusan Masalah, Tujuan dan Manfaat Penelitian, Pembatasan Masalah, dan Sistematika Penulisan dari penelitian ini.
2. Bab II Kajian Literatur  
Membahas tentang Kinerja Siswa, Data Mining, Pengelompokan, Algoritma *Fuzzy C-Means Clustering*, Algoritma ROCK, Penelitian yang Berkaitan, dan Kerangka Konsep pada penelitian ini.
3. Bab III Metodologi Penelitian  
Berisi tentang Analisis Masalah, Data yang Digunakan, Alat-Alat Penelitian, Rancangan Penelitian, Contoh Perhitungan Algoritma ROCK, dan Contoh Perhitungan Algoritma FCM pada penelitian ini.
4. Bab IV Hasil dan Pembahasan  
Berisi tentang Uji Coba, Hasil yang Diperoleh dari setiap pengujian, dan Pembahasan dari penelitian yang dilakukan.
5. Bab V Kesimpulan dan Saran  
Menyimpulkan hasil atas penelitian yang telah dilakukan dan saran yang diberikan bagi penelitian selanjutnya.