

BAB I

PENDAHULUAN

1.1.Latar Belakang

Perkembangan teknologi Indonesia yang semakin pesat telah menandakan adanya perubahan transisi pada era kehidupan manusia yang baru. Hal ini tidak dapat dipungkiri bahwa kehidupan manusia telah memasuki di era digital 5.0 yang memberikan dampak dan efek dengan bertujuan untuk mempermudah segala pekerjaan manusia. Dalam bidang teknologi memberikan pengaruh terhadap sektor bidang di kehidupan manusia baik dari segi ekonomi, sosial, kesehatan hingga di pemerintahan. Terdapat banyak perubahan yang terjadi akibat dari perkembangan teknologi salah satunya terjadinya pemanfaatan sistem digital yang awalnya setiap data di kolektifkan secara konvensional saat ini data yang dikumpulkan dari instansi sudah berbasis komputer. Melalui pemanfaatan teknologi dan informasi tiap harinya setiap instansi dapat menghasilkan hingga triliunan byte data. Data-data tersebut berupa tulisan, gambar, suara, angka, dan lain-lain. Dari sekian banyaknya data yang di kumpulkan setiap waktunya hal ini mengarah kepada pemanfaatan Big data yang dapat mengolah serta menganalisis data-data tersebut menjadi lebih mudah dibaca dan dicari apabila data tersebut di butuhkan (Mustari *et al.*, 2024).

Data mining atau yang sering disebut KDD (*Knowledge Discovery in Database*) adalah proses yang mencakup pengumpulan dan pemanfaatan riwayat data untuk menemukan pola dan hubungan dalam kumpulan data yang besar. Penambangan data menghasilkan informasi yang dapat digunakan untuk membuat keputusan di masa mendatang. Salah satu teknik yang dikenal dalam penambangan data adalah pengelompokan atau pengelompokan data (*clustering*) (Lestari *et al.*, 2024).

Data adalah informasi dan angka yang dikumpulkan untuk digunakan dalam berbagai jenis penelitian dan analisis sebagai dukungan untuk pandangan atau pendapat (Setiawan & Ramdhani, 2024).

Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar (Sasmita & Rahmadayanti, 2024).

Data Mining memfasilitasi pengguna untuk mengungkap informasi yang mungkin tidak akan terlihat oleh mereka saat melakukan eksplorasi data dalam basis data. Data Mining adalah salah satu komponen dalam proses *Knowledge Discovery in Databases* (KDD), yang melibatkan serangkaian langkah, termasuk pemilihan data, pra-pemrosesan, transformasi data, ekstraksi informasi, dan evaluasi hasil. KDD sering disebut juga sebagai penemuan pengetahuan dalam basis data. Data Mining memiliki lima fungsi utama, termasuk pengelompokan (*clustering*), klasifikasi (*classification*), asosiasi (*association*), urutan (*sequencing*) dan peramalan (*forecasting*) (Liantoni *et al.*, 2024).

Clustering adalah salah satu teknik yang termasuk dalam fungsi data mining. Algoritma clustering digunakan untuk mengelompokkan sejumlah data ke dalam

kelompok-kelompok tertentu, yang disebut cluster. Data mining, yang sering disebut sebagai *Knowledge Discovery in Database (KDD)*, adalah proses yang mencakup pengumpulan dan penggunaan data historis untuk mengidentifikasi pola, hubungan, atau keteraturan dalam data berukuran besar (Pujiono *et al.*, 2024).

Algoritma *clustering* berupaya memisahkan kumpulan informasi yang ada menjadi kelompok-kelompok yang homogen atau sejenis. Tingkat kesamaan data di dalam suatu kelompok akan menghasilkan nilai yang semakin besar, sementara perbedaan antar kelompok akan menghasilkan nilai yang semakin kecil (Lestari *et al.*, 2024).

K-means merupakan salah satu metode pengelompokan data non-hierarki yang mempartisi data yang ada ke dalam bentuk dua atau lebih kelompok (Surapati & Jannah, 2024).

Algoritma *K-Means* merupakan bagian dari clustering data mining, dimana algoritma *K-Means* dapat dipergunakan untuk pembentukan kelompok baru dari data. Pembentukan kelompok baru dari data pada algoritma *K-Means* dengan proses pembentukan *cluster* pada proses yang dilakukan (Faran & Triayudi, 2024).

Penelitian sebelumnya dilakukan oleh Iin, Rizal Fadila, Ade Rizki Rinaldi, Fathurrohman yang berjudul “Penerapan Data Mining Dalam Mengelompokan Jumlah UMKM

Berdasarkan Kabupaten Kota Menggunakan *K-Means Clustering*” bertujuan adalah untuk mengelompokan kabupaten atau kota di provinsi Jawa Barat berdasarkan tahun dan jumlah usaha mikro kecil menengah (UMKM) dengan menerapkan teknik data mining menggunakan Algoritma *K-Means clustering*. mengevaluasi hasil dari proses clustering dengan nilai *index davies bouldin*. Implikasi hasil dari penelitian ini memiliki potensi yang baik bagi para pelaku

UMKM untuk mengetahui analisis wilayah mereka masuk kedalam kategori kelompok yang mana, terhadap masyarakat umum bisa menambah wawasan atau penegetahuan baru mengenai hasil pengelompokan, dan bagi pemerintahan baik daerah atau pusat bisa memberikan kontribusi dalam pengambilan keputusan terhadap sektor UMKM. Penelitian ini bisa menjadi referensi pemahaman terhadap proses penerapan data mining dengan algoritma *K-Means* menggunakan *software Rapidminer 10.2*. *RapidMiner* adalah perangkat lunak terbuka yang dapat digunakan untuk melakukan analisis pada data mining, text mining dan analisis prediktif (Fadila & Rinaldi, 2024).

Penelitian selanjutnya bertujuan untuk mengetahui tingkat pemahaman materi mata kuliah inti pada program studi Informatika Universitas Dehasen yang diterima oleh mahasiswa, dimana peneliti menyebarkan kuisioner kepada mahasiswa untuk menentukan tingkat pemahaman materi mata kuliah inti menjadi 4 kelompok yaitu sangat baik, baik, cukup baik dan kurang baik. Dari hasil tersebut di implementasikan ke aplikasi *Rapidminer*.

Dari uraian di atas, maka penulis mengangkat penelitian berupa tesis judul **"Data Mining Dengan Metode *K-Means* Dalam Mengukur Tingkat Pemahaman Materi Mata Kuliah Inti Dan Penilaian Mahasiswa (Studi Kasus Di Prodi Informatika Universitas Dehasen Bengkulu)"**.

1.2.Perumusan Masalah

Clustering data menjadi alternative yang menarik pada penelitian dalam mengukur tingkat pemahaman materi mata kuliah inti dan penilaian mahasiswa. Penelitian yang dilakukan saat ini menimbulkan beberapa perumusan masalah pokok antara lain:

1. Bagaimana penerapan metode *K-Means* dalam mengukur tingkat pemahaman materi mata kuliah inti dan penilaian mahasiswa?
2. Bagaimana penerapan aplikasi *RapidMiner* dalam mengukur tingkat pemahaman materi mata kuliah inti dan penilaian mahasiswa
3. Bagaimana tingkat pemahaman materi mata kuliah inti dan penilaian mahasiswa terhadap Prodi Informatika Universitas Dehasen Bengkulu?

1.3. Batasan Masalah

Agar tidak terjadi penyimpangan dan lebih terarah dalam laporan penelitian ini, maka penelitian diberi batasan-batasan sebagai berikut :

1. Data yang diolah hanya pada Program Studi Informatika Fakultas Ilmu Komputer Dehasen Bengkulu
2. Dalam menguji kebenaran hasil, penulis menggunakan *software RapidMiner*

1.4. Tujuan Penelitian

Tujuan penelitian yang ingin diperoleh penulis dalam penelitian ini adalah sebagai berikut:

1. Untuk menerapkan metode *K-Means* dalam Mengukur Tingkat Pemahaman materi mata kuliah inti dan penilaian mahasiswa
2. Menerapkan aplikasi *RapidMiner* dalam mengukur Tingkat pemahaman materi mata kuliah inti dan penilaian mahasiswa
3. Untuk mengetahui tingkat pemahaman materi dan penilaian mahasiswa terhadap Prodi Informatika Universitas Dehasen Bengkulu?

1.5. Manfaat Penelitian

Penelitian ini bertujuan untuk dapat memberikan manfaat yang dapat berguna bagi yang membutuhkan penelitian sejenis, adapun manfaat penelitian yang akan dilakukan adalah sebagai berikut:

1. Memberi referensi terhadap pengembangan data mining
2. Pengembangan data mining untuk peningkatan kualitas pengajaran dosen
3. Membantu pihak akademik dalam mengevaluasi proses kegiatan mengajar dosen

1.6. Sistematika Penulisan

Dalam menyusun karya tulis ilmiah ini, agar dalam pembahasan terfokus pada pokok permasalahan dan tidak melebar kemasalah yang lain, maka penulis membuat sistematika penulisan karya tulis ilmiah sebagai berikut :

BAB I PENDAHULUAN

Pada bab ini dijelaskan tentang latar belakang masalah, perumusan masalah, batasan masalah, tujuan penelitian dan manfaat penelitian serta sistematika penulisan.

BAB II LANDASAN TEORI

Pada bab ini membahas tentang landasan teori dan penerapannya tentang data mining, mata kuliah inti, algoritma *K-Means* serta teori-teori lain yang mendukung penelitian ini.

BAB III METODOLOGI PENELITIAN

Pada bab ini dijabarkan secara detail tentang kerangka penelitian serta metodologi-metodologi yang digunakan dalam meng-*cluster* pemahaman materi mata kuliah inti Prodi Informatika

BAB IV ANALISA DAN PERANCANGAN

Pada bab ini akan dibahas tentang analisa pemahaman materi mata kuliah inti Prodi Informatika dengan algoritma *K-Means*, bagaimana meng-*cluster* data mata kuliah secara teoritis dengan cara menggunakan Aplikasi *RapidMiner*

BAB V IMPLEMENTASI DAN HASIL

Pada bab ini akan membahas mengenai hasil dari penilaian pemahaman mahasiswa terhadap mata kuliah dengan algoritma *K-Means* secara teoritis dan hasil penilaian dengan algoritma *K-Means* berdasarkan aplikasi *RapidMiner*.

BAB VI PENUTUP

Pada bab ini memuat kesimpulan dan saran yang didapat setelah melakukan penelitian dan pengujian dengan menggunakan algoritma *K-Means*.

