

BAB I

PENDAHULUAN

1.1. LATAR BELAKANG

Penentuan jurusan akan berdampak terhadap kegiatan akademik selanjutnya dan mempengaruhi pemilihan bidang ilmu atau studi bagi siswa-siswi yang ingin melanjutkan ke perguruan tinggi nantinya. Penentuan jurusan yang dilakukan selama ini mempunyai banyak kelemahan, antara lain berdasarkan keinginan siswa tanpa melihat latar belakang nilai akademisnya. Sehingga jurusan yang dipilih terkadang menjadi masalah bagi siswa di kemudian hari, sebagai contoh nilai akademik yang tidak maksimal, pemilihan program studi saat melanjutkan ke jenjang perguruan tinggi yang terkendala akibat jurusan SMA yang tidak sesuai, dan lain-lain.

Berdasarkan hasil wawancara dengan wakil kesiswaan Dwi Wahyuningsih, M.Pd.,kons mengatakan, siswa-siswi SMA Negeri 2 Kota Jambi cenderung memilih jurusan berdasarkan karena minat, dan keinginan orang tua. Beberapa di antaranya sudah memperhitungkan potensi yang ada pada diri mereka, maka komitmen untuk belajar dibidang itu tidak akan berjalan lancar, padahal jurusan yang dia pilih itu tidak sesuai kemampuannya. Harapannya pihak sekolah bisa melihat persentase keakuratannya untuk penentuan jurusan Perguruan Tinggi Negeri pada siswa-siswi SMA Negeri 2 Kota Jambi, jika metode ini berhasil dan persentasenya tinggi, pihak sekolah akan menggunakan kembali metode ini untuk

merekomendasikan jurusan Perguruan Tinggi Negeri untuk siswa-siswi selanjutnya.

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. tujuan utama *data mining* adalah untuk menemukan, menggali, atau menambang pengetahuan dari data atau informasi yang kita miliki (Yuda Septian Nugroho, 2014 : 2-3).

Teknologi *clustering* data merupakan suatu teknik yang menunjukkan persamaan karakteristik dalam suatu kelompok sehingga akan menghasilkan informasi yang bermanfaat. Algoritma *clustering* data sudah banyak dipergunakan diberbagai bidang misalnya untuk proses pengolahan citra, data mining proses pengambilan keputusan, pengenalan pola, maupun dalam bidang bioinformatika (Menurut Debatty, Thibault., et.Al, 2014). Ada beberapa algoritma yang untuk dapat melakukan proses *clustering* pada suatu dataset dalam jumlah yang banyak. Pada penelitian ini, peneliti akan menggunakan metode algoritma *K-Means* dalam menentukan jumlah cluster terbaik. *K-Means* merupakan algoritma yang sangat banyak dipergunakan karena efektif dan efisien. Ini dikarenakan *K-means* sangat mudah dipelajari dan dari segi waktu proses komputasinya relatif singkat (Menurut Kaur, K., Dhaliwal, D.S. & Vohra, K.R. 2013).

Dari permasalahan yang di temui pada SMA Negeri 2 Kota Jambi, serta berdasarkan hasil penelitian sebelumnya, maka penulis akan melakukan pengelompokan jurusan perguruan tinggi bagi siswa-siswi SMA Negeri 2 Kota

Jambi dengan penerapan *data mining*. Agar penelitian ini menyelesaikan masalah penentuan jurusan ke perguruan tinggi untuk siswa-siswi SMA Negeri 2 Kota Jambi yang dituangkan dalam bentuk tugas akhir skripsi yang di beri judul **“PENERAPAN *CLUSTERING* DATA MINING UNTUK REKOMENDASI PENENTUAN JURUSAN PERGURUAN TINGGI PADA SISWA SMA NEGERI 2 KOTA JAMBI MENGGUNAKAN METODE *K-MEANS*”**. Dengan adanya penelitian ini diharapkan tingkat persentase kelulusan dalam menentukan jurusan Perguruan Tinggi Negeri pada siswa-siswi SMA Negeri 2 Kota Jambi lebih tinggi.

1.2. RUMUSAN MASALAH

1. Bagaimana menerapkan *k-means clustering* untuk merekomendasi penentuan memilih jurusan di perguruan tinggi setelah lulus SMA ?
2. Bagaimana mengevaluasi dan menganalisis dari hasil *clustering* menggunakan algoritma *k-means* untuk rekomendasi penentuan jurusan perguruan tinggi pada siswa-siswi SMA Negeri 2 Kota Jambi ?

1.3 BATASAN MASALAH

Untuk menghindari pembahasan yang meluas, maka penulis hanya membatasi pembahasan permasalahan hanya pada.

1. Objek yang menjadi sasaran penelitian adalah siswa/i kelas XII di SMA NEGERI 2 Kota Jambi dengan jumlah 210 siswa.

2. Atribut yang di gunakan yaitu : Hobi, Minat, Bakat, Sifat, nilai rata-rata mata pelajaran untuk jurusan IPA, dan nilai rata-rata mata pelajaran untuk jurusan IPS
3. Analisis menggunakan teknik *data mining clustering* dengan metode *k-means*.
4. Data yang digunakan bersumber dari SMA NEGERI 2 Kota Jambi
5. Pengujian hasil analisis menggunakan *tools WEKA* dan *RapidMiner*.
6. Format data yang digunakan adalah Arff.

1.4. TUJUAN DAN MANFAAT PENELITIAN

1.4.1. Tujuan Penelitian

Penelitian ini bertujuan untuk:

1. Menerapkan teknik *data mining* dengan *K-Means Clustering* dalam merekomendasikan jurusan pada Perguruan Tinggi Negeri pada SMA NEGERI 2 Kota Jambi.
2. Mengevaluasi hasil perhitungan *Clustering* dengan *Algoritma K-Means* pada siswa-siswi SMA Negeri 2 Kota Jambi.

1.4.2. Manfaat Penelitian

1. Dapat memprediksi siswa SMA NEGERI 2 KOTA JAMBI dalam menentukan jurusan pada Perguruan Tinggi Negeri mana yang direkomendasikan pada berdasarkan nilai.

2. Pihak SMA NEGERI 2 KOTA JAMBI dapat menentukan strategi untuk meningkatkan kualitas dan kuantitas pada pembelajaran untuk tahun-tahun berikutnya.
3. Penulis dapat menambah ilmu dan wawasan baru mengenai *Clustering*., menentukan jurusan pada Perguruan Tinggi Negeri pada siswa SMA NEGERI 2 KOTA JAMBI.
4. Dapat digunakan sebagai acuan untuk melakukan penelitian berikutnya.

1.5 SISTEMATIKA PENULISAN

Adapun sistematika penulisan tugas akhir ini, penulis menguraikan dalam beberapa bab yaitu :

BAB I : PENDAHULUAN

Pada bab ini akan diuraikan tentang latar belakang masalah, rumusan masalah, batasan masalah, tujuan dan manfaat penulisan, dan sistematika penulisan.

BAB II : LANDASAN TEORI

Pada bab landasan teori ini membahas tentang teori-teori dan pendapat para ahli yang berhubungan dengan permasalahan yang dianalisis. Teori-teori yang digunakan antara lain mengenai *data mining*, *Clustering*., prediksi, lama masa studi, *K-Means*, dan *tools data mining*.

BAB III : METODOLOGI PENELITIAN

Pada bab ini menjelaskan tentang kerangka kerja penelitian, metode pengumpulan data, metode *Clustering* serta alat bantu yang digunakan pada penelitian ini.

BAB IV : ANALISIS

Pada bab ini dilakukan perhitungan analisis menggunakan metode *Clustering K-Means* terhadap data-data nilai siswa yang tersedia.

BAB V : HASIL ANALISIS DAN VISUALISASI

Pada bab ini akan ditampilkan hasil dari analisis dan bentuk visualisasi analisis dari tools *WEKA* dan *RapidMiner* yang digunakan.

BAB VI : PENUTUP

Bab ini yang berisikan kesimpulan-kesimpulan yang diambil dari hasil analisis serta saran-saran yang mencakup keseluruhan dari hasil penelitian.