

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Sesuai yang berlaku di Indonesia para siswa di SMA diwajibkan memilih jurusan yang sesuai dengan minat, bakat dan kemampuan yang dimiliki oleh setiap siswa. Pemilihan jurusan ini amat sangat berpengaruh untuk kelanjutan masa depan baik yang akan menempuh pendidikan perguruan tinggi maupun yang akan mencari pekerjaan (Putri et al., 2022). faktor-faktor yang mempengaruhi siswa dalam pemilihan jurusan dibagi menjadi dua kelompok. Pertama, faktor dari dalam diri siswa yang meliputi minat, kepribadian, dan citra/konsep diri. Sedangkan faktor dari luar diri remaja meliputi orangtua, teman sebaya, lingkungan sosial ekonomi budaya, dan saran tes bakat minat (Hartati et al., 2018).

Dalam menentukan bidang ilmu yang cocok untuk siswa diperlukan kriteria tepat, pada umumnya dalam menentukan atau melihat keahlian dan peminatan siswa terhadap satu bidang studi adalah dengan melihat nilai akademik setiap mata pelajaran dan minat bakat siswa, dan diolah untuk mendapatkan hasil jurusan yang terbaik untuk siswa.

Pengolahan data secara manual akan memakan waktu yang sangat lama, karena banyaknya data siswa yang ada di sekolah, Pada era sekarang perkembangan teknologi khususnya informasi dalam pengolahan data membuat masyarakat dituntut untuk menimba pengetahuan dan meningkatkan kemampuan dibidang teknologi dalam pengolahan data (Nasir, 2021).

Begitu juga dengan pihak sekolah dapat memanfaatkan perkembangan ilmu *data science*, perkembangan *data science* berkembang sangat pesat seperti halnya kebutuhan manusia. Dalam hal ini *data science* memiliki fungsi signifikan terutama dalam hal pengolahan data. Pengolahan data berdasarkan *data science* dapat digunakan untuk mengklastering siswa (Primanda et al., 2021). pengolahan data untuk mengklaster siswa juga dapat diterapkan pada kasus pemilihan bidang studi atau bidang ilmu yang tepat untuk siswa dalam menempuh pendidikan tinggi

selanjutnya, salah satu cabang ilmu yang dapat membantu untuk mengolah data agar dapat diklaster adalah *Data Mining Clustering*.

Data mining merupakan proses dalam menemukan hubungan yang berarti, pola dan tren dengan memeriksa data berukuran besar dalam suatu penyimpanan dengan menggunakan teknologi pengenalan pola, misalnya statistik dan matematika. Dari beberapa pengertian sebelumnya, dapat disimpulkan bahwa Data Mining merupakan serangkaian proses dalam pencarian pola, hubungan, penggalian nilai tambah dari data dan informasi yang berukuran besar berupa pengetahuan dengan tujuan menemukan hubungan dan menyederhanakan data agar diperoleh informasi yang dapat dipahami dan bermanfaat dengan bantuan ilmu statistik dan matematika.

Dalam Data Mining, umumnya metode penelitian berkaitan dengan klasifikasi, clustering, dan Asosiasi. Dan dalam kasus ini akan memanfaatkan metode *Clustering*, metode *Clustering* merupakan metode pengelompokan data, observasi, atau kasus menjadi kelas objek-objek yang serupa. Sedangkan *cluster* didefinisikan sebagai kumpulan data yang sama satu sama lain, dan tidak sama dengan data di lain *cluster* (Abdurrahman, 2016). dan salah satu algoritma yang digunakan untuk menentukan bidang studi siswa SMA adalah algoritma *K-Means Clustering*, Algoritma *K-means* merupakan salah satu algoritma *clustering* (pengelompokan). *K-means clustering* merupakan metode *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 kecil (Abdurrahman, 2016).

Pada penelitian ini juga dapat sejalan dengan penerapan *Active Learning in Machine Learning*, dalam beberapa kutipan jurnal dan literatur *Active Learning in Machine Learning* bertujuan untuk membangun secara optimal model *Supervised Learning*, optimasi model adalah salah satu cara untuk dapat menghasilkan kinerja terbaik dengan biaya yang minimum dengan klasifikasi manual, *Active Learning* adalah subbidang dari *Machine Learning* sebagai *Learner* (Pembelajar) untuk diizinkan meminta contoh yang paling informatif untuk melatih ulang

model dari membuat pilihan secara acak, atas dasar pendekatan ini, dengan jumlah pemilihan sample yang sama, yaitu kinerja dari *Active Learner* melebihi cara acak lainnya (Nguyen & Patrick, 2014).

Untuk fungsi dan kelebihan dari penerapan *Active Learning* ini sendiri adalah *Active Learning* (AL) dapat mencapai akurasi sebanding dengan *Passive Learning* dengan lebih sedikit label, hal ini sudah di amati dalam banyak pengaplikasian praktis dalam beberapa dekade terakhir (Hanneke, 2014), dengan kelebihan yang dimiliki AL itu sendiri peneliti dapat memudahkan proses penelitian ini dengan lebih cepat dan menghemat biaya karena sudah menggunakan AL untuk melakukan perbaikan terhadap hasil dari model *Clustering*, dalam kasus ini hanya dibutuhkan *Human Annotator* (manusia) untuk memperbaiki akurasi *model*. Proses ini dilakukan dengan memilih dan mengidentifikasi data yang paling sulit diklasifikasikan atau memiliki tingkat *incertitude* (ketidakpastian) tinggi, dan meminta annotator untuk memberikan label pada data tersebut. AL dapat digabungkan dengan algoritma *Clustering Data Mining* itu sendiri, dan juga AL tidak memiliki rumus tunggal karena AL adalah suatu pendekatan yang digunakan untuk pembelajaran mesin.

Pemilihan bidang studi atau program keahlian bagi siswa SMA terutama di SMAS Muhammadiyah 2 Medan menjadi hal yang penting, walaupun SMA adalah sekolah yang sudah menempatkan siswa pada minat jurusan tertentu, akan tetapi faktanya masih banyak ditemukan siswa menempuh jurusan yang sangat berbeda pada saat di perguruan tinggi. Lalu pemilihan bidang studi masih banyak mengalami kendala, seperti tidak tersedianya informasi yang cukup dan sesuai dengan minat dan bakat siswa, atau kurangnya asesmen yang akurat mengenai kemampuan siswa. Untuk mengatasi masalah tersebut, solusi yang dapat dilakukan adalah dengan menggunakan teknik *data mining* dan *machine learning* untuk membantu menentukan bidang studi yang sesuai.

Dalam hal ini, algoritma *K-Means Clustering* dapat digunakan untuk melakukan pengelompokan siswa berdasarkan minat dan bakat mereka. Namun, hasil pengelompokan algoritma *K-Means Clustering* masih memiliki tingkat akurasi yang belum maksimal. Oleh karena itu, penerapan AL dapat menjadi solusi untuk meningkatkan akurasi hasil pengelompokan. *Active learning*

meminta masukan atau validasi dari siswa untuk memvalidasi hasil clustering dan memperbaiki akurasi klasifikasi bidang studi yang diminati.

Berdasarkan latar belakang masalah diatas, maka peneliti tertarik melakukan penelitian dengan judul “Implementasi Algoritma *K-Means Clustering* Dengan Pendekatan *Active Learning* Pada Siswa SMA Untuk Menentukan Jurusan ke Perguruan Tinggi”. Penelitian ini diharapkan dapat menyelesaikan permasalahan siswa yang tidak mengetahui dan tidak dapat menemukan minat bakat dan kemampuan diri sendiri, dan pihak sekolah lebih tepat dalam mengarahkan siswa pada jurusan yang cocok untuk siswa untuk menempuh Pendidikan tinggi.

1.2. Identifikasi Masalah

Berdasarkan dari latar belakang masalah dapat diidentifikasi beberapa hal sebagai berikut :

1. Banyaknya jurusan di Perguruan Tinggi yang belum diketahui siswa: Siswa tidak dikenalkan kepada jurusan yang tepat sesuai minat bakatnya, pengenalan jurusan secara umum hanya akan menambah rasa bingung terhadap keputusan jurusan apa yang harus mereka ambil, hal ini akan lebih terarah jika sudah dilakukan penentuan jurusan dengan Model *K-Means Clustering*.
2. Siswa tidak menemukan minat bakat yang tepat diri sendiri: Dalam penelitian ini didapat suatu masalah siswa yang sulit menemukan minat bakat diri sendiri, hal ini di faktori oleh kurangnya kepercayaan diri dan keyakinan terhadap kemampuan soal bakat diri sendiri, dan siswa tidak memiliki pengetahuan tentang bagaimana cara menemukan dan melihat minat bakat siswa dengan cara yang tepat.
3. Siswa tidak mengetahui cara menemukan minat diri sendiri: Model akan lebih membantu dalam menemukan minat bakat siswa dengan melihat beberapa atribut atau variabel yaitu berupa hasil Angket Minat/Bakat siswa dan Rata-rata Nilai Rapot Siswa.

1.3. Ruang Lingkup Masalah

Penelitian ini melibatkan penggunaan Model Algoritma *K-Means Clustering* dengan pendekatan *Active Learning* pada siswa di SMAS Muhammadiyah 2 Medan untuk menentukan jurusan ke Perguruan Tinggi berdasarkan hasil *clustering* Model dengan atribut atau variabel Hasil Kuesioner Minat/Bakat siswa, dan Rata-rata Nilai Rapot Siswa .

1.4. Batasan Masalah

Pembatasan masalah dibuat untuk menghindari dan mengantisipasi adanya pelebaran terhadap pembahasan pokok suatu masalah, agar penelitian yang dilakukan dapat lebih terarah dan mudah dalam melakukan pembahasan. Untuk Batasan masalah yang digunakan dalam penelitian ini adalah sebagai berikut :

1. Algoritma *K-Means Clustering* dengan pendekatan *Active Learning* digunakan untuk menentukan dan menempatkan siswa SMAS Muhammadiyah 2 Medan sesuai dengan rumpun ilmu dan jurusan yang tepat untuk diambil di Perguruan Tinggi.
2. Penentuan *cluster* siswa SMAS Muhammadiyah 2 Medan diolah dengan Model dan data didapat dari hasil kuesioner minat/bakat siswa dan Rata-rata Nilai Rapot Siswa.
3. *Cluster* dibagi kedalam 12 *cluster*, mewakili berbagai rumpun ilmu yang berbeda sesuai dengan sumber dari Direktorat Jenderal Pendidikan Tinggi tentang 12 Koding Rumpun ilmu yang ada di Indonesia, yaitu Rumpun Matematika dan Ilmu Pengetahuan Alam, Ilmu Tanaman, Ilmu Hewani, Ilmu Kedokteran, Ilmu Kesehatan, Teknik, Ilmu Bahasa, Ilmu Ekonomi, Ilmu Sosial Humaniora, Agama dan Filsafat, Seni Desain dan Media, dan Ilmu Pendidikan.
4. Penelitian ini dilakukan dengan menggunakan 50% data siswa kelas 12 pada semua jurusan di SMAS Muhammadiyah 2 Medan
5. Perhitungan evaluasi model *k-means clustering* menggunakan *silhouette score*.

1.5. Rumusan Masalah

Berdasarkan latar belakang yang sudah di uraikan sebelumnya, permasalahan yang ada pada penelitian ini adalah sebagai berikut:

1. Bagaimana mengimplementasikan Model Algoritma *K-Means Clustering* dengan pendekatan *Active Learning*?
2. Bagaimana menentukan hasil implementasi Algoritma *K-Means Clustering* dengan pendekatan *Active Learning* untuk menentukan jurusan ke Perguruan Tinggi pada data siswa?

1.6. Tujuan Penelitian

Berdasarkan rumusan masalah yang sudah diuraikan diatas, Tujuan penelitian ini adalah sebagai berikut :

1. Mengimplementasikan Model Algoritma *K-Means Clustering* dengan pendekatan *Active Learning*
2. Menentukan hasil implementasi Algoritma *K-Means Clustering* dengan pendekatan *Active Learning* untuk menentukan jurusan ke Perguruan Tinggi pada data siswa

1.7. Manfaat Penelitian

Adapun manfaat yang diharapkan pada penelitian ini yaitu:

1. Bagi Peneliti, Peneliti dapat menambah pengetahuan baru tentang bagaimana menentukan jurusan siswa SMAS Muhammadiyah 2 Medan ke Perguruan Tinggi dengan melihat hasil dari model yang dibuat berdasarkan atribut atau variabel Hasil Kuesioner Minat/Bakat siswa dan Rata-rata Nilai Rapot Siswa dan dapat mengimplementasikan Algoritma *K-Means Clustering* dengan pendekatan *Active Learning* sebagai bentuk memperkuat keterampilan di bidang *Data Mining* dan *Machine Learning*
2. Bagi Pembaca, Pembaca dapat menambah informasi tentang bagaimana menentukan jurusan siswa SMAS Muhammadiyah 2 Medan ke Perguruan Tinggi dengan mengimplementasikan Algoritma *K-Means*

Clustering dengan pendekatan *Active Learning* pada data siswa berupa Hasil Angket Minat/Bakat siswa dan Rata-rata Nilai Rapot Siswa.

3. Bagi Akademik, Hasil penelitian ini dapat menambah informasi bagi pihak sekolah SMAS Muhammadiyah 2 Medan mengenai bagaimana cara menentukan jurusan siswa SMAS Muhammadiyah 2 Medan ke Perguruan Tinggi dengan mengimplementasikan Algoritma *K-Means Clustering* dengan pendekatan *Active Learning* pada data siswa berupa hasil Angket Minat/Bakat siswa dan Rata-rata Nilai Rapot Siswa. Selain itu, pihak sekolah juga dapat menerapkan penelitian ini agar siswa lulusan SMAS Muhammadiyah 2 Medan kedepannya dapat memaksimalkan kemampuan minat/bakat dalam menempuh pendidikan perguruan tinggi, hal itu akan menaikkan kualitas sekolah.
4. Bagi Siswa, Siswa SMAS Muhammadiyah 2 Medan akan lebih mudah melihat dan menentukan minat/bakat pada diri sendiri, dengan mengacu kepada hasil dari model Algoritma *K-Means Clustering* dengan pendekatan *Active Learning*, hasil model akan lebih tepat mengarahkan siswa dalam memilih jurusan ke perguruan tinggi negeri.