

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

SMP Islam Temayang Merupakan Sekolah Menengah Pertama yang berada pada Kecamatan Temayang Kabupaten Bojonegoro. Siswa yang berada pada SMP Islam Temayang terbagi dalam beberapa kelas. Dalam Satu angkatan bisa di bagi dalam Beberapa kelas Yaitu kelas A,B,C dan D Karena Banyaknya Siswa yang berada pada SMP Islam Temayang, maka dalam penempatan siswa di setiap kelas juga perlu diperhatikan dalam meningkatkan mutu pendidikan dan prestasi siswa SMP Islam Temayang. Peningkatan mutu pendidikan merupakan prioritas kebijakan umum pembangunan di Indonesia yang harus terus diperhatikan(Dewi, Aryni, and Umaidah 2022). Menurut undang- undang nomor 20 tahun 2003, tujuan Pendidikan nasional ialah untuk menaikkan keterampilan terhadap peserta didik supaya menjadi manusia yang beriman dan juga bertakwa kepada Tuhan Yang Maha Esa, berakhlak mulia, sehat, berilmu, cakap, kreatif, mandiri, dan juga diharapkan menjadi penduduk negara yang demokratis serta memiliki rasa bertanggung jawab. Bersumber pada tujuan pendidikan tersebut sehingga indikator kualitas sekolah dapat dilihat dari prestasi belajar maupun akademik siswa.

Prestasi ialah perubahan dalam perihal kecakapan dalam bertindak, maupun keahlian yang bisa berkembang sepanjang beberapa waktu serta tidak diakibatkan proses perkembangan, namun karena adanya suasana belajar(Prasetyawan and Gatra 2022). Lingkungan teman sekelas sangat berpengaruh terhadap perkembangan prestasi akademik siswa, jika hubungannya negatif maka akan berdampak negatif bagi siswa, dan juga terjadi sebaliknya (Alvian et al. 2022). Namun, terdapat banyak sekolah yang menerapkan penggolongan dalam pembentukan kelas agar mempermudah guru dalam kegiatan belajar mengajar, dimana siswa yang pintar akan sekelas dengan siswa yang pintar dan juga sebaliknya, dalam hal ini menyebabkan beberapa siswa belum mencapai prestasi akademik yang maksimal. Untuk menghindari

masalah di atas, beberapa sekolah menerapkan pengacakan dalam pembentukan kelas.

SMP Islam Temayang salah satu sekolah di Kecamatan Temayang yang menerapkan pengacakan pada pembentukan kelasnya, dimana siswa yang tergolong pintar atau mendapat predikat A akan disebar ke semua kelas seangkatan untuk membantu siswa lain yang memiliki masalah dalam belajar, penyebaran tersebut harus merata agar efektif, siswa yang memiliki prestasi akademik dan non akademik bagus biasanya adalah siswa yang aktif didalam kelas dan merupakan siswa dengan peringkat 5 besar dikelasnya masing-masing (Alvian et al. 2022). Predikat atau kategori itu sendiri didapatkan dari perhitungan aspek nilai akademik dan non akademik. Dikarenakan belum adanya metode khusus yang digunakan untuk mengklasifikasikan siswa berdasarkan prestasinya dan banyak kemiripan data, menyebabkan sering terjadi salah pengklasifikasian dan penyebaran siswa pintar yang tidak merata sehingga hasilnya menjadi kurang maksimal. Oleh karena itu dibutuhkan metode klasifikasi data mining yang tepat dan akurat dalam penentuan siswa berprestasi pada SMP Islam Temayang.

Klasifikasi merupakan pembelajaran terawasi yang memprediksi label dari suatu kelas. Klasifikasi digunakan untuk menemukan model yang menjelaskan atau membedakan kelas data yang dapat memperkirakan kelas dari suatu data yang kelasnya belum diketahui (Prasetyawan and Gatra 2022). Metode klasifikasi seperti *C4.5*, *Decision Tree*, *Naïve Bayes*, *K-Nearest Neighbor* banyak digunakan untuk klasifikasi siswa berprestasi. Algoritma *C4.5* merupakan teknik data mining yang dapat melakukan klasifikasi serta prediksi dengan cara mengolah beberapa *variable* yang akan digunakan. Algoritma *C4.5* merupakan algoritma yang dikembangkan dari algoritma *ID3*. (Husein and Hutaeruk 2022) algoritma *C4.5* ini merupakan algoritma turunan dari algoritma *ID3* dengan beragam peningkatan. Beberapa peningkatan ini diantaranya adalah, penanganan pada atribut-atribut numerik, *missing value* dan *noise* pada dataset, dan aturan-aturan yang dihasilkan dari model pohon yang terbentuk (Iriadi and Nuraeni 2016).

Decision Tree merupakan salah satu metode klasifikasi yang mudah dipahami (Lakshmi, Indumathi, and Ravi 2016). Bentuk representasi dari metode ini dengan menggunakan hierarki pohon (*tree*) yang di dalamnya terdapat sebuah *node* dan *leaf* (Permana, Ainiyah, and Holle 2021). Kelebihan dari metode ini bersifat fleksibel yang dapat meningkatkan tingkat akurasi yang dihasilkan. Selain itu, *Decision Tree* dapat mempersingkat atribut pada setiap *node* tanpa banyak menguras akurasi yang dihasilkan. Disamping itu, metode ini terjadi *overlap* jika penggunaan kelas dan kriterianya banyak (Bisri and Wahono 2015).

Naive Bayes merupakan sebuah metode pengklasifikasian probabilistik sederhana yang dapat menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi serta kombinasi nilai dari dataset yang ada (Yos, Km, and Mulia 2016). Definisi lain mengatakan *Naive Bayes* merupakan pengklasifikasian dengan metode probabilitas serta statistik yang dikemukakan oleh ilmuwan Inggris bernama Thomas Bayes, yaitu memprediksi peluang yang ada di masa depan berdasarkan pengalaman dimasa sebelumnya (Bustami 2019). Keuntungan penggunaan *Naive Bayes* adalah bahwa metode ini hanya membutuhkan jumlah data pelatihan (*Training Data*) yang kecil dalam menentukan estimasi parameter yang diperlukan pada proses pengklasifikasian. *Naive Bayes* sering bekerja lebih baik pada kebanyakan situasi dunia nyata yang kompleks dari pada yang diharapkan.

Metode klasifikasi dengan algoritma *K-Nearest Neighbor* (KNN) merupakan sesuatu metode yang memakai algoritma supervised, dimana hasil dari sampel uji yang baru diklasifikasikan bersumber pada kebanyakan dari jenis pada *K-Nearest Neighbor* (KNN) yang bertujuan untuk mengklasifikasi objek baru bersumber pada atribut serta training samples (Alvian et al. 2022). Algoritma *k-nearest neighbor* (k-NN) memiliki beberapa kelebihan yaitu bahwa k-NN tangguh terhadap training data yang noise (Ratri, Wali, and Wahyuni 2020) K-NN juga dapat menghasilkan data yang kuat dan efektif jika digunakan pada data dengan jumlah yang cukup besar (Nugroho and Wijana 2018), yang mana metode ini tepat untuk pengklasifikasian siswa berprestasi

sehingga dapat digunakan untuk penerapan metode klasifikasi *K-Nearest Neighbor* (KNN) untuk mengklasifikasikan siswa berprestasi pada SMP Isalm Temayang.

Berdasarkan penelitian sebelumnya, yaitu Prediksi Prestasi Siswa SMP Nurul Jadid Menggunakan Algoritma C4.5 oleh (kamil malik, mohammad faid, 2016), membahas tentang bagaimana proses klasifikasi prestasi siswa dengan menggunakan data sebanyak 1006 dan 4 atribut. Berdasarkan hasil penelitian, algoritma C4.5 menghasilkan akurasi sebesar 72.4652%. dan Pada penelitian lainnya Prediksi Prestasi Mahasiswa dengan Jalur Siswa Berprestasi (PSB) oleh (Miratama 2016) Membahas tentang klasifikasi siswa berprestasi menggunakan metode naïve bayes dengan 77 data dan 7 atribut menghasilkan tingkat akurasi sebesar 51,69%.

Berdasarkan latar belakang yang telah dijelaskan sebelumnya, maka pada penelitian yang dilakukan kali ini akan berjudul “Klasifikasi Siswa Berprestasi Berbasis Kriteria Akademik Dan Non Akademik Menggunakan Algoritma *K-Nearest Neighbor*”. Dengan menggunakan atribut aspek akademik seperti nilai mata pelajaran, serta nilai aspek non akademik seperti absensi dan kegiatan ekstrakurikuler serta ditambahkan variable berupa prestasi yang menaikan nilai tambah bagi siswa. Dengan tujuan untuk mengukur tingkat akurasi dan presisi algoritma *K-Nearest Neighbor* (KNN) dalam menentukan klasifikasi siswa berprestasi. Pada penelitian ini data yang digunakan sebanyak 81 data dan dilakukan split data 7 banding 3 sehingga diperoleh hasil 56 data sebagai data training dan 25 data sebagai data testing dan diperoleh akurasi sebesar 87,50%.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang di atas, maka permasalahan yang akan dibahas dalam penelitian ini adalah:

1. Bagaimana seleksi siswa berprestasi menggunakan atribut aspek akademik dan non akademik sehingga diperoleh katagori siswa berprestasi ?
2. Berapa tingkat akurasi dan presisi dari algoritma *K-Nearest Neighbor* (KNN) jika diterapkan dalam proses klasifikasi siswa berprestasi ?

1.3 Batasan Masalah

Agar permasalahan tidak menyimpang dari tujuan penelitian, maka berikut beberapa batasan yang perlu diperhatikan, yaitu:

1. Ruang lingkup penelitian ini dilakukan di SMP Islam Temayang
2. Data yang digunakan adalah siswa kelas VII tahun 2022 di SMP Islam Temayang.
3. Fitur atau Input yang digunakan pada penelitian ini adalah Aspek Akademik dan Non Akademik.
4. Output yang dihasilkan adalah katagori siswa berprestasi dan tidak berprestasi.
5. Perhitungan jarak menggunakan rumus *euclidean distance*.

1.4 Tujuan Penelitian

Adapun tujuan dari penelititan ini adalah :

1. Untuk mengkalsifikasi siswa berprestasi menggunakan atribut aspek akademik dan non akademik sehingga diperoleh siswa berprestasi yang nantinya disebar pada setiap kelas secara merata untuk meningkatkan potensi belajar siswa yang lain
2. Mengetahui tingkat akurasi serta presisi dari penggunaan algoritma *K-Nearest Neighbor* (KNN) dengan metode pengukuran jarak *euclidean distance* dalam proses klasifikasi siswa berprestasi pada SMP Islam Temayang.

1.5 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah:

1. Sebagai metode alternatif klasifikasi siswa berprestasi di SMP Islam Temayang
2. Hasil dari penelitian ini diharapkan dapat memberi kesempatan untuk menambah wawasan dan pengetahuan dalam bidang klasifikasi siswa berprestasi.
3. Mempermudah guru menyebarkan siswa berprestasi ke kelas seangkatan secara merata.