

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Pendidikan merupakan komponen penting dalam pembangunan bangsa dan pembentukan kualitas sumber daya manusia. Di Indonesia, pendidikan dasar selama sembilan tahun merupakan kewajiban. Pendidikan adalah usaha sadar dan terencana untuk mewujudkan suasana belajar serta proses pembelajaran agar peserta didik secara aktif mengembangkan potensi dirinya, baik dalam aspek spiritual, emosional, intelektual, maupun keterampilan sosial yang berguna bagi diri sendiri, masyarakat, bangsa, dan negara (Undang-Undang Nomor 20 Tahun 2003).

Dalam implementasinya, sekolah menjadi wadah utama pelaksanaan pendidikan formal, di mana guru dan siswa berinteraksi langsung dalam proses belajar mengajar. Prestasi belajar siswa menjadi indikator penting untuk menilai keberhasilan pendidikan. Penilaian prestasi ini umumnya berdasarkan pada pencapaian akademik dan partisipasi siswa dalam proses pembelajaran, seperti kehadiran dan keterlibatan aktif di kelas (Rahayu et al., 2024). Namun, tidak semua siswa bisa menunjukkan kinerja akademik yang optimal. Hal ini menandakan bahwa prestasi belajar dipengaruhi oleh banyak faktor yang kompleks dan saling berinteraksi, sehingga diperlukan pendekatan sistematis untuk menganalisisnya.

Salah satu contoh nyata dari permasalahan tersebut dapat dilihat di SMPN 2 Kesamben, Kabupaten Jombang. Sebuah sekolah negeri dengan akreditasi A dan jumlah siswa mencapai 576 orang. Meski memiliki sumber daya yang baik, sekolah ini masih menghadapi tantangan dalam meningkatkan prestasi belajar sebagian siswanya. Oleh karena itu, penting dilakukan analisis mendalam untuk mengidentifikasi faktor-faktor yang berkontribusi terhadap prestasi akademik dan memprediksi kecenderungan kinerja belajar siswa di masa depan.

Seiring perkembangan teknologi informasi, metode data mining menjadi alat yang sangat potensial dalam dunia pendidikan. Data mining memungkinkan penggalian informasi tersembunyi dari data yang besar untuk mendukung pengambilan keputusan. Salah satu metode yang umum digunakan dalam prediksi

prestasi belajar adalah algoritma K-Nearest Neighbor (KNN), yang memanfaatkan kedekatan antar data untuk mengklasifikasikan atau memprediksi nilai target.

Metode K-Nearest Neighbor (KNN) dipilih karena kemudahan penggunaannya, dan akurasinya yang cukup baik dalam memprediksi data seperti prestasi belajar siswa. KNN tidak memerlukan aturan atau rumus statistik yang kompleks, sehingga sangat sesuai untuk diterapkan dalam dunia pendidikan, terutama ketika data yang digunakan terdiri dari berbagai jenis, seperti nilai angka, kehadiran, atau aktivitas ekstrakurikuler (Sakti dan Daulay 2024). Selain itu, KNN beroperasi dengan membandingkan kemiripan antara data baru dan data lama, sehingga semakin mirip datanya, semakin tinggi kemungkinan hasil prediksinya akurat. Menurut (Steven 2022), metode ini juga terbukti efektif dalam klasifikasi data siswa karena mampu mengidentifikasi pola dari data numerik dan kategori secara bersamaan.

Penelitian sebelumnya yang dilakukan oleh (Widaningsih 2022) dengan judul "Penerapan Data Mining untuk Memprediksi Siswa Berprestasi dengan Menggunakan Algoritma K Nearest Neighbor". Tujuan dari penelitian ini adalah untuk membantu pihak sekolah, khususnya SMK Tunas Sinar Mandiri Cianjur, dalam memprediksi siswa berprestasi di kelas XII menggunakan teknik data mining dengan algoritma k-Nearest Neighbor (kNN). Penilaian tidak hanya didasarkan pada aspek akademik, tetapi juga mempertimbangkan faktor non-akademik seperti keaktifan, kepribadian, kehadiran, dan jenis kelamin. Dari hasil pengujian, model ini mampu mencapai akurasi sebesar 86% dalam memprediksi siswa berprestasi.

Namun, terdapat kesenjangan dalam penelitian ini yang dapat dimanfaatkan untuk penelitian selanjutnya yaitu penelitian ini hanya mencakup nilai produktif, keaktifan, kepribadian, presensi, dan jenis kelamin, tanpa mempertimbangkan faktor lain yang mungkin relevan seperti variabel nilai ekstrakurikuler secara spesifik sebagai prediktor. Variabel ini penting karena keterlibatan dalam kegiatan ekstrakurikuler sering kali mencerminkan kepemimpinan, tanggung jawab, dan pengembangan karakter siswa, yang merupakan indikator penting dari prestasi non-akademik.

Penelitian ini bertujuan untuk mengembangkan model prediksi prestasi belajar siswa dengan menerapkan metode K-Nearest Neighbor (KNN), serta mengevaluasi dampak dari berbagai faktor akademik dan non-akademik terhadap prestasi tersebut

dengan mengintegrasikan nilai ekstrakurikuler secara spesifik sebagai variabel prediktor tambahan yang dapat meningkatkan akurasi model dan memberikan gambaran yang lebih komprehensif mengenai karakteristik siswa yang berprestasi.

Dengan memanfaatkan pendekatan data mining, diharapkan hasil penelitian ini dapat memberikan bantuan kepada pihak sekolah, khususnya di SMPN 2 Kesamben Kabupaten Jombang, dalam mengidentifikasi dan mempersiapkan siswa yang memiliki potensi unggul, baik dalam aspek akademik maupun pengembangan karakter.

## **1.2 Rumusan Masalah**

Berdasarkan identifikasi masalah yang telah diuraikan diatas, maka didapatkan rumusan masalah dalam penelitian ini yaitu :

- 1) Bagaimana cara penerapan algoritma K-Nearest Neighbor (KNN) untuk memprediksi prestasi belajar siswa di SMPN 2 Kesamben dengan menggunakan data akademik dan non-akademik?
- 2) Bagaimana dampak penambahan variabel nilai ekstrakurikuler terhadap tingkat akurasi model prediksi prestasi belajar siswa yang memanfaatkan algoritma KNN?

## **1.3 Tujuan Penelitian**

Berdasarkan rumusan masalah tersebut, tujuan yang ingin dicapai dalam penelitian ini sebagai berikut :

- 1) Menerapkan algoritma K-Nearest Neighbor (KNN) untuk membangun model prediksi prestasi belajar siswa berbasis data akademik dan non-akademik di SMPN 2 Kesamben.
- 2) Menambahkan variabel nilai ekstrakurikuler sebagai atribut prediktor guna meningkatkan kualitas dan akurasi hasil prediksi.

## **1.4 Batasan Masalah**

Ruang lingkup penelitian ini cukup luas sehingga dibutuhkan batasan masalah agar penelitian lebih terarah. Maka batasan masalah pada penelitian ini adalah sebagai berikut :

- 1) Penelitian ini menggunakan data prestasi belajar siswa yang tersedia di SMPN 2 Kesamben. Sehingga hasil penelitian ini mungkin terbatas, pada data yang ada diambil dari siswa kelas VII, VIII, dan kelas IX semester ganjil pada tahun akademik 2024-2025.
- 2) Penelitian ini akan terbatas pada penggunaan metode K-Nearest Neighbor (K-NN) sebagai algoritma untuk analisis data, tanpa membandingkan dengan metode lain seperti Naive Bayes atau yang lainnya.
- 3) Fokus penelitian akan terbatas pada variabel tertentu yang dapat diukur, seperti nilai rapor, kehadiran, dan nilai ekstrakurikuler tanpa mempertimbangkan faktor eksternal lainnya seperti kondisi sosial dan ekonomi siswa.
- 4) Data yang digunakan dalam penelitian ini akan diambil dari sistem informasi akademik sekolah dan tidak mencakup data dari sumber lain seperti survei atau wawancara dengan siswa dan orang tua.

### **1.5 Manfaat Penelitian**

Berdasarkan uraian penelitian ini dapat memberikan manfaat kepada beberapa pihak yaitu :

- 1) Penelitian ini dapat membantu untuk memperdalam pemahaman tentang konsep data mining metode K-Nearest Neighbor, serta penerapannya dalam bidang pendidikan.
- 2) Memenuhi salah satu syarat kelulusan Mahasiswa Teknik Informatika Universitas Darul `Ulum Jombang.
- 3) Penelitian ini dapat membantu pihak sekolah dalam memprediksi prestasi belajar siswa secara lebih cepat dan akurat sehingga memudahkan dalam keputusan yang lebih tepat dalam meningkatkan kualitas pendidikan.
- 4) Dapat menggunakan hasil prediksi sebagai bahan evaluasi dan strategi pembelajaran yang lebih efektif.
- 5) Penelitian ini dapat membantu guru untuk mengidentifikasi siswa yang berpotensi mengalami kesulitan belajar dan memberikan bantuan dalam proses pembelajaran.

## **1.6 Sistematika Penulisan**

Penulisan tugas akhir ini memiliki sistematika penulisan, adapun uraian masing- masing bab sebagai berikut :

### **1) BAB I : PENDAHULUAN**

Dalam bab ini, meliputi latar belakang penelitian, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan yang digunakan dalam penyusunan Tugas Akhir ini.

### **2) BAB II : LANDASAN TEORI**

Pada bab ini, terdiri dari penelitian terdahulu yang dapat mendukung penelitian ini dan uraian teori relevan mengenai penelitian ini.

### **3) BAB III : METODE PENELITIAN**

Bab ini terdiri dari prosedur penelitian meliputi tahapan mulai dari identifikasi masalah, pengumpulan data, seleksi data, transformasi data, preprocessing data, pembagian dataset, pengujian model dan menampilkan hasil rapidminer. Teknik pengumpulan data yaitu observasi dan wawancara, dan ruang lingkup penelitian dalam penelitian ini.

### **4) BAB IV : HASIL DAN PEMBAHASAN**

Untuk bab ini akan menjelaskan hasil penerapan metode penelitian dan pembahasan menggunakan metode K- Nearest Neighbor dengan pengujian menggunakan Rapidminer dan secara manual (Excel).

### **5) BAB V : PENUTUP**

Bab ini terdiri dari kesimpulan berdasarkan hasil penelitian dan saran penulis kepada berbagai pihak untuk penelitian selanjutnya.

### **6) DAFTAR PUSTAKA**

Bagian ini menjelaskan tentang kumpulan atau daftar sumber referensi yang digunakan dalam penyusunan penelitian ini.