

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

Pendidikan tidak terlepas dari kegiatan pembelajaran. Prestasi belajar bagi siswa sangat penting karena prestasi belajar merupakan salah satu gambaran tingkat keberhasilan dari kegiatan selama mengikuti pelajaran. Salah satu tujuan dalam proses pembelajaran adalah meraih suatu prestasi dalam belajar. Prestasi belajar merupakan hasil belajar yang diperoleh oleh siswa setelah mengikuti proses belajar mengajar yang dilakukan oleh guru. Prestasi belajar siswa dapat diketahui setelah diadakan evaluasi. Hasil dari evaluasi dapat memperlihatkan tentang tinggi atau rendahnya prestasi belajar siswa (Sardiman, 2011).

Sekolah Menengah Pertama AL-Gazali Jember adalah salah satu sekolah yang mempersiapkan peserta didik khususnya siswa kelas sembilan untuk siap menghadapi Ujian Nasional Berbasis Komputer (UNBK). Salah satu upaya yang dilakukan SMP Al-Gazali adalah melakukan kerja sama dengan pihak luar seperti Ganesha dalam mengadakan *Tryout*. Hasil-hasil *Tryout* dari pihak luar kemudian dijadikan bahan pertimbangan untuk menganalisis kesiapan siswa dalam menghadapi UNBK.

*Data Mining* merupakan proses pencarian pola-pola yang menarik dan tersembunyi (*hidden pattern*) dari suatu kumpulan data yang berukuran besar yang tersimpan dalam sebuah basis data, *data warehouse*, atau penyimpanan data lainnya (Tan, et al, 2006). *Educational Data Mining* adalah cara pendekatan yang muncul berkaitan dengan pengembangan metode untuk menjelajahi jenis yang unik dari data yang berasal dari lingkungan pendidikan dan menggunakan metode tersebut untuk lebih memahami para siswa. Salah satu teknik *Data Mining* yang bisa diterapkan pada lingkungan pendidikan adalah klasifikasi. Klasifikasi merupakan bentuk analisis data yang menghasilkan sebuah model untuk menjelaskan kelas data penting. Beberapa tahun terakhir, telah terjadi peningkatan minat dalam penggunaan *Data Mining* di bidang pendidikan. *Educational Data Mining* (EDM) berkaitan dengan bagaimana mengembangkan metode dan

menganalisis konten pendidikan untuk memungkinkan pemahaman yang lebih baik terhadap kinerja siswa. Hal ini juga penting untuk meningkatkan proses belajar mengajar. Ada berbagai metode EDM dan algoritma yang digunakan untuk menemukan pola dan hubungan yang meliputi prediksi, *clustering* dan asosiasi. Tujuan dari prediksi adalah untuk mengembangkan model yang dapat menyimpulkan satu aspek data (diprediksi parameter) dari beberapa kombinasi data lain (parameter *prediktor*). Pola yang ditemukan dengan metode *Data Mining* dari data pendidikan dapat digunakan untuk meningkatkan pengambilan keputusan dalam hal mengidentifikasi siswa yang berisiko, penurunan tingkat *drop-out* mahasiswa, meningkatkan keberhasilan siswa, dan meningkatkan hasil belajar siswa. Pola yang diekstrak juga dapat menemukan perilaku belajar siswa dan karakteristik yang mempengaruhi kinerja akademis mereka.

Pada penelitian terdahulu yaitu pada penelitian dengan judul *Data Mining Analysis on Student's Academic Performance through Exploration of Student's Background and Social Activities* menggunakan metode *Data Mining* diantaranya *Naïve Bayesian*, *Multilayer Perceptron*, *Decision Tree J48* dan *Random Forest* dalam melakukan klasifikasi siswa untuk mengukur tingkat performa (Kie, et al, 2018). *K-Nearest neighbor* merupakan salah satu metode klasifikasi dalam *Data Mining*. Beberapa penelitian juga menggunakan metode ini dalam hal *Education Data Mining*. Pada penelitian dengan judul Sistem Pendukung Keputusan Untuk Menentukan Status Prestasi Siswa Menggunakan Metode *K- Nearest Neighbor* menggunakan metode KNN dalam mengklasifikasi data siswa dan guna menentukan prestasi hasil belajar (Irawan, et al, 2013). Dari uraian di atas penulis tertarik untuk melakukan penelitian dengan judul “Analisis Pemodelan Metode *K-Nearest Neighbor* Terhadap Klasifikasi Kesiapan Siswa Kelas IX SMP Al Gazali Dalam Menghadapi UNBK”.

## 1.2. Rumusan Masalah

Dalam penelitian ini penulis menitikberatkan permasalahan pada;

1. Berapa tingkat akurasi tertinggi pada hasil klasifikasi metode K-NN terhadap data siswa?

2. Berapa tingkat presisi tertinggi pada hasil klasifikasi metode K-NN terhadap data siswa?

### 1.3. Tujuan

Dari hasil analisis penelitian ini hal yang dituju antara lain;

1. Mengetahui tingkat akurasi tertinggi dari hasil klasifikasi metode K-NN terhadap data siswa.
2. Mengetahui tingkat presisi tertinggi dari hasil klasifikasi metode K-NN terhadap data siswa.

### 1.4. Manfaat

Dari hasil penelitian ini dapat dimanfaatkan terhadap;

1. Bahan pertimbangan sekolah untuk mengetahui kesiapan siswa dalam menghadapi ujian nasional berbasis komputer.
2. Pengembangan *Education Data Mining* dengan atribut yang berbeda.
3. Membantu guru dalam mempersiapkan siswa menghadapi ujian nasional berbasis komputer.

### 1.5. Batasan Masalah

Dalam penelitian ini penulis memiliki batasan pembahasan diantaranya;

1. Data penelitian yang digunakan yaitu data kelas sembilan dari SMP Al-Gazali Jember tahun ajaran 2019-2020.
2. Pada penerapan metode K-NN, pengukuran vektor jarak menggunakan *Euclidean Distance* dengan nilai  $k = 3, 5, 7$  dan  $9$ .
3. Skenario uji menggunakan *K Fold Cross Validation* dengan nilai  $K = 2, 4, 5, 7$  dan  $10$
4. Pengujian performansi yang dilakukan adalah pengukuran tingkat akurasi dan presisi.
5. Untuk mengatasi *Imbalance Data* menggunakan metode SMOTE.