

# BAB I PENDAHULUAN

## I.1 Latar Belakang

Pendidikan pada hakikatnya merupakan suatu usaha terpolara untuk menciptakan suasana belajar dan proses pembelajaran agar pelajar secara aktif menghasilkan potensi diri untuk memiliki kecerdasan dan keterampilan yang dibutuhkan. Pendidikan juga merupakan alur untuk mendapatkan pengetahuan dalam proses pembelajaran serta untuk mendapatkan pengetahuan yang luas dalam melatih pemahaman serta keterampilan diri dengan cara melakukan latihan soal (Fathoni, 2019). Naskah soal adalah sekumpulan soal yang telah dikelompokkan dengan kecocokan materi dan jenis dari ujian yang diberikan untuk pelajar berdasarkan soal – soal ujian yang telah dibuat sebelumnya (Listiowarni & Puspa, 2020).

Naskah soal digunakan sebagai evaluasi pengajar untuk mengetahui sejauh mana kemampuan pelajar menyelesaikan soal berdasarkan tingkat kesulitan (F. T. Dinata & Rosyana, 2021). Isi dari naskah soal biasanya berupa kisi-kisi soal yang menyediakan contoh soal dan jumlah persentase dari tingkat kesulitan soal. Naskah soal perlu disesuaikan kembali untuk pengajar dapat memberikan berbagai soal yang sesuai pengetahuan pelajar berdasarkan level kognitif *Bloom's Taxonomy* (Listiowarni & Puspa, 2020).

*Bloom's Taxonomy* merupakan klasifikasi yang dapat digunakan untuk memprediksi kemampuan pelajar dalam belajar sebagai penilaian dari kegiatan pembelajaran (Sujoko & Darmawan, 2013). Mengapa menggunakan *Bloom's Taxonomy*? ini merupakan salah satu kerangka dasar ranah kognitif untuk klasifikasi yang telah digunakan setengah abad untuk penyusunan unsur pendidikan, penyusunan soal, dan kurikulum (Gunawan & Paluti, 2017) dan *Bloom's Taxonomy* juga merupakan model taksonomi yang memiliki tujuan sebagai dasar pengembangan kurikulum dalam sistem pendidikan di Indonesia (Minarto, 2020).

Untuk melakukan klasifikasi soal berdasarkan *Bloom's Taxonomy* merupakan tugas yang sulit jika dilakukan secara manual (Kusuma dkk., 2017), yang berdampak pada proses pengolahan data akan membutuhkan waktu yang cukup

lama terhadap jumlah data kapasitas besar (Aninditya dkk, 2019). Dengan melakukan klasifikasi soal secara otomatis dengan bantuan *machine learning* berdasarkan *Bloom's Taxonomy*, penelitian ini bertujuan untuk mempermudah pengajar dapat memberikan penilaian proses pembelajaran dan mengevaluasi untuk mempertimbangkan tolak ukur hasil kemampuan pelajar terhadap tingkat pengetahuan pelajar dari latihan soal yang diberikan (Adams, 2015).

Oleh sebab itu, penulis mengangkat topik ini untuk melakukan klasifikasi tipe soal berdasarkan level kognitif berdasarkan *Revised Bloom's Taxonomy* yang terfokus tiga tingkat, yaitu mengingat (C1), memahami (C2), mengaplikasikan (C3) yang bertujuan untuk membantu proses klasifikasi tipe soal mata pelajaran Sejarah Indonesia tingkat SMA berdasarkan level kognitif RBT dalam membuat naskah soal. Terdapat salah satu metode klasifikasi yang dapat digunakan yaitu metode algoritma KNN atau *K-Nearest Neighbour* dengan pendekatan jarak *Manhattan*. Secara singkat, *K-Nearest Neighbour* merupakan metode *machine learning* untuk klasifikasi dengan menghitung jarak dengan nilai  $k$  tetangga terdekat pada data *training* (Ramadhani & Murti, 2018) dan KNN merupakan algoritma klasifikasi yang memiliki performa yang cukup baik pada penelitian (Listiowarni & Puspa, 2020). Penggunaan jarak *Manhattan* mampu meningkatkan hasil akurasi klasifikasi serta pendekatan jarak ini merupakan pengukuran *similarity* paling tepat untuk *approval project* yang mewakili kasus yang menggunakan bilangan asli atau data kuantitatif. (Latifah, 2015).

Terdapat penelitian terdahulu yang menjadi rujukan pada penelitian lanjutan ini untuk melakukan klasifikasi soal berdasarkan level kognitif *Revised Bloom's Taxonomy* menggunakan algoritma KNN. Pada penelitian (Dinata dkk., 2020) yang menguji perbandingan algoritma KNN dengan pendekatan jarak *Manhattan* dan *Euclidean* dalam menghitung jarak pada pengklasifikasian transportasi bus yang diperoleh nilai akurasi 84% dengan jarak *Manhattan* lebih besar dibandingkan jarak *Euclidean* sebesar 81,96% . Selanjutnya pada penelitian (Pantiwati & Permana, 2017) yang menguji kualitas butir soal berdasarkan PISA dan *Revised Bloom's Taxonomy*, untuk hasil yang didapatkan sebagian besar masih belum memenuhi standar berdasarkan PISA dan *Revised Bloom's Taxonomy*. Hasil akhir dari penelitian ini akan menghasilkan persentase tingkat

akurasi dari klasifikasi soal Sejarah Indonesia tingkat SMA berdasarkan level kognitif RBT menggunakan algoritma *K-Nearest Neighbour* dengan pendekatan jarak *Manhattan* dan penggunaan jarak *Manhattan* digunakan pada penelitian ini dipercaya memiliki nilai akurasi yang lebih tinggi (Dinata dkk., 2020)

## **I.2 Perumusan Masalah**

Berdasarkan latar belakang penelitian ini, rumusan masalah yang akan dilakukan adalah sebagai berikut:

1. Bagaimana mengatasi ketidakseimbang pada dataset soal mata pelajaran Sejarah Indonesia berdasarkan level kognitif C1 – C3 *Revised Bloom's Taxonomy*?
2. Bagaimana melakukan klasifikasi soal secara otomatis mata pelajaran Sejarah Indonesia berdasarkan level kognitif C1 – C3 *Revised Bloom's Taxonomy* dengan algoritma KNN?
3. Bagaimana hasil tingkat akurasi klasifikasi pada dataset soal mata pelajaran Sejarah Indonesia dengan metode KNN?

## **I.3 Tujuan Penelitian**

Penelitian ini memiliki tujuan di antaranya:

1. Menjelaskan solusi cara mengatasi data yang tidak seimbang pada proses klasifikasi tipe soal Sejarah Indonesia berdasarkan level kognitif C1 – C3 *Revised Bloom's Taxonomy*,
2. Menjelaskan bagaimana penerapan algoritma KNN untuk melakukan klasifikasi secara otomatis pada soal Sejarah Indonesia berdasarkan level kognitif *Revised Bloom's Taxonomy*,
3. Mengetahui hasil tingkat akurasi dalam proses klasifikasi tipe soal Sejarah Indonesia menggunakan algoritma KNN.

## **I.4 Batasan Penelitian**

Adapun batasan dalam penelitian ini dilakukan sebagai berikut:

1. Penggunaan algoritma *K-Nearest Neighbour* dengan menggunakan pendekatan jarak *Manhattan* untuk klasifikasi soal,

2. *Tools* yang digunakan adalah *Python*,
3. Mengklasifikasi soal berdasarkan level kognitif *Revised Bloom's Taxonomy* yang berfokus pada C1 – C3,
4. Data yang digunakan berupa kumpulan soal mata pelajaran Sejarah Indonesia tingkat SMA dengan jumlah 246 soal yang didapat dari internet,
5. Metode SMOTE digunakan untuk mengatasi *imbalance data* atau ketidakseimbangan data,
6. Berfokus pada tipe soal teks pilihan ganda mata pelajaran Sejarah Indonesia.

## **I.5 Manfaat Penelitian**

Manfaat penelitian ini adalah:

1. Untuk pengajar, diharapkan dapat memberikan soal latihan sesuai tingkat kesulitan sebagai bahan evaluasi penilaian dan mengetahui tingkat kemampuan pelajar dalam mengerjakan soal – soal yang diberikan.
2. Untuk pelajar, diharapkan dapat membantu sebagai bahan evaluasi diri berdasarkan tingkat kemampuan diri dalam mengerjakan berbagai tipe soal yang diberikan oleh pengajar.
3. Untuk penelitian, dapat menjelaskan bagaimana hasil dari persentase prediksi dan tingkat keakuratan dalam melakukan klasifikasi tipe soal berdasarkan tingkat kesulitan sebagai bahan rujukan dalam melakukan penelitian lanjutan.
4. Untuk penulis, dapat menerapkan penggunaan algoritma *K-Nearest Neighbour* dalam proses klasifikasi soal serta mengetahui hasil dari prediksi persentase dan tingkat keakuratannya pada proses klasifikasi tipe soal berdasarkan tingkat kesulitan *Revised Bloom's Taxonomy*.