

BAB I PENDAHULUAN

I.1 Latar Belakang

Pendidikan secara umum bertujuan untuk mengembangkan kepribadian diri manusia karena terdapat proses pengajaran ilmu pengetahuan dan keterampilan (*skills*). Sesuai dengan tujuan tersebut, disebutkan pada UU. No. 20 Tahun 2003 pasal 3 tentang tujuan pendidikan yakni mengembangkan potensi peserta didik agar menjadi manusia yang beriman dan bertakwa kepada Tuhan yang Maha Esa, berakhlak mulia, sehat, berilmu, cakap, kreatif, mandiri serta menjadi warga negara yang demokratis serta bertanggung jawab (Hakim, 2016).

Sebagai upaya untuk menjamin siswa memahami materi yang diajarkan, para guru mengadakan uji kompetensi yang dibagi dalam beberapa termin setiap semesternya. Ujian merupakan bentuk evaluasi proses belajar dalam rangka mengukur taraf pencapaian keahlian, karakter serta intelegensi siswa, sehingga menjadikan ujian sebagai tahapan penting dalam proses belajar mengajar (Sidiq & Kurniadi, 2021).

Higher order thinking skills (HOTS) atau keterampilan berpikir tingkat tinggi merupakan bagian dari *Bloom's Taxonomy* hasil revisi yang berupa kata kerja operasional yang terdiri dari *analyze* (C4), *evaluate* (C5) dan *create* (C6) yang dapat digunakan dalam penyusunan soal. Guru harus memiliki pengetahuan dan keahlian guna menunjang pekerjaan, hingga dapat mengembangkan keterampilan berpikir tingkat tinggi peserta didik tingkat SMA (Aydin & Yilmaz, 2010).

Guru menyusun bank soal salah satunya pada platform *e-Learning* untuk menghimpun pertanyaan yang akan diujikan untuk siswa. Penyusunan soal membutuhkan pemilahan hingga hanya soal tingkat tinggi yang dapat mengukur tingkat intelektual siswa dan mencakup keterampilan berpikir yang akan terpilih (Pantiwati dkk., 2017). Mata pelajaran sejarah termasuk tipe pelajaran dengan banyak teori dan bacaan. Penggunaan *e-Learning* pada mata pelajaran ini penting karena mata pelajaran sejarah dapat melandasi pendidikan intelektual dan kedisiplinan siswa (Nurcholif dkk., 2021). Meskipun begitu, penyusunan bank soal *e-Learning* terutama pada mata pelajaran sejarah memerlukan evaluasi ulang

dikarenakan banyaknya soal yang belum tersaring sesuai dengan capaian kompetensinya. Proses penyaringan melibatkan algoritma identifikasi, salah satunya yang sering digunakan dalam peningkatan kualitas soal yaitu *Bloom's Taxonomy* (Mohammed & Omar, 2018).

Bloom's Taxonomy dalam ranah kognitif merupakan salah satu kerangka dasar untuk melakukan kategorisasi tujuan-tujuan pendidikan, penyusunan tes, dan kurikulum (Gunawan & Paluti, 2017). Secara umum klasifikasi soal berdasarkan *Bloom's Taxonomy* dilakukan secara manual, namun cara tersebut dinilai relatif lama untuk mengolah data dalam skala besar (Aninditya et al., 2019). Oleh karena itu, penulis mengajukan solusi untuk mengadakan klasifikasi soal secara otomatis dengan menggunakan salah satu metode *machine learning*. Metode *machine learning* diketahui dapat dipakai untuk memproses soal ujian dengan menggunakan NLP (Kusuma et al., 2016).

Pada beberapa tahun terakhir, terdapat beberapa masalah dalam proses klasifikasi yang dihadapi yaitu terkait dengan capaian akurasi berdasarkan kesesuaian informasi yang didapatkan (Elsalamony & Elsayad, 2018). Dalam menyelesaikan permasalahan klasifikasi tersebut, penggunaan metode atau teknik bertujuan untuk mempermudah proses klasifikasi.

Metode yang akan digunakan pada penelitian ini yaitu *Stochastic Gradient Descent* (SGD). SGD menggunakan *gradien stochastic* yang meminimalkan fungsi kerugian yang dipilih dengan fungsi linear (Dong & Chen, 2009). Algoritma ini mendekati gradien yang benar dengan mempertimbangkan satu sampel pada suatu waktu, dan secara bersamaan memperbarui model berdasarkan gradien fungsi kerugian (Ruder, 2016). Meskipun memiliki aktivitas *update* varian yang tinggi, SGD dipilih karena dapat diatur dalam pengurangan *learning rate* sehingga menghasilkan *loss function* dengan titik minimum (Khotimah et al., 2019). SGD telah membuktikan pengoptimalan yang efisien pada data yang masif, hal ini karena SGD sering kali dapat memperoleh gradien yang *noisy* tetapi tidak bias dengan melakukan *subsampling* secara acak pada kumpulan data yang besar (Mandt dkk., 2017).

Berdasarkan uraian di atas, penelitian ini akan mengusulkan melakukan model dalam teknik klasifikasi pada *data mining* bank soal. Penelitian ini akan menghasilkan klasifikasi tipe soal berdasarkan tingkat kognitif, tingkat akurasi, dan prediksi guna penentuan klasifikasi soal sejarah SMA tipe C4 hingga C6.

I.2 Perumusan Masalah

Berdasarkan latar belakang di atas, dihasilkan rumusan masalah penelitian yaitu sebagai berikut:

1. Bagaimana cara menangani *imbalance data* pada *dataset* soal-soal mata pelajaran sejarah pada jenjang SMA berdasarkan tingkat kognitif C4-C6?
2. Bagaimana hasil implementasi algoritma SGD dalam mengklasifikasikan soal-soal mata pelajaran sejarah pada jenjang SMA berdasarkan tingkat kognitif C4-C6?
3. Bagaimana tingkat akurasi yang dihasilkan dari klasifikasi menggunakan algoritma SGD dalam mengklasifikasikan soal-soal mata pelajaran sejarah pada jenjang SMA berdasarkan tingkat kognitif C4-C6?
4. Bagaimana perolehan hasil akurasi antara *dataset* yang menggunakan metode *oversampling* dengan *dataset* tanpa menggunakan metode *oversampling*?

I.3 Tujuan Penelitian

Penelitian ini bertujuan untuk:

1. Menjelaskan cara menangani *imbalance data* pada *dataset* soal-soal mata pelajaran sejarah pada jenjang SMA berdasarkan tingkat kognitif C4-C6.
2. Mengetahui hasil implementasi algoritma SGD dalam mengklasifikasikan soal-soal mata pelajaran sejarah pada jenjang SMA berdasarkan tingkat kognitif C4-C6.
3. Mengetahui tingkat akurasi yang dihasilkan dari klasifikasi menggunakan algoritma SGD dalam mengklasifikasikan soal-soal mata pelajaran sejarah pada jenjang SMA berdasarkan tingkat kognitif C4-C6.
4. Mengetahui perolehan hasil akurasi antara *dataset* yang menggunakan metode *oversampling* dengan *dataset* tanpa menggunakan metode *oversampling*.

I.4 Batasan Penelitian

Batasan dalam penelitian ini antara lain:

1. Data yang digunakan adalah soal mata pelajaran sejarah SMA,
2. Jumlah soal adalah sebanyak 58 butir yang dihimpun dari bank soal di internet,
3. Berfokus pada tipe soal yang diberikan untuk siswa SMA yaitu soal berbahasa Indonesia pilihan ganda berbentuk teks,
4. Soal diklasifikasikan berdasarkan tingkat kognitif C4-C6 (Bloom's Taxonomy),
5. Algoritma yang digunakan adalah SGD,
6. Metode yang digunakan untuk mengatasi *imbalance data* adalah SMOTE,
7. *Tools* yang digunakan adalah *Phyton*.

I.5 Manfaat Penelitian

Manfaat penelitian ini dikategorikan menjadi dua yaitu manfaat teoritis dan manfaat praktis. Manfaat teoritis sebagai kontribusi dalam referensi mengenai klasifikasi pertanyaan khususnya pada bank soal mata pelajaran sejarah SMA berdasarkan tingkat kognitif C4-C6 menggunakan algoritma SGD. Manfaat praktis penelitian ini memiliki dampak bagi berbagai pihak yang dapat diimplementasikan dalam kehidupan, di antaranya:

1. Bagi Pengajar
Penelitian ini diharapkan mampu membantu pengajar untuk mengevaluasi kembali soal mata pelajaran sejarah pada jenjang SMA sesuai tingkat kognitif C4-C6 menggunakan hasil penelitian bermodel algoritma SGD.
2. Bagi Siswa
Penelitian ini diharapkan mampu membantu siswa untuk mendapatkan soal mata pelajaran sejarah pada jenjang SMA sesuai tingkat kognitif C4-C6 sehingga dapat meningkatkan efisiensi pembelajaran.
3. Bagi Penulis
Penelitian ini diharapkan mampu membantu penulis untuk dapat mengimplementasikan performa algoritma SGD untuk mengklasifikasi soal mata pelajaran sejarah pada jenjang SMA, serta mengetahui tingkat akurasi soal berdasarkan tingkat kognitif C4-C6.