

BAB I

PENDAHULUAN

Bab pendahuluan ini membahas mengenai latar belakang masalah, rumusan masalah, tujuan dan manfaat, ruang lingkup, serta sistematika penulisan yang digunakan dalam skripsi berjudul *Automatic Short Answer Grading* pada soal uraian berbahasa Indonesia menggunakan SBERT dan metode *K-Means Clustering*.

1.1 Latar Belakang

Pendidikan merupakan fondasi utama dalam pembangunan manusia dan masyarakat. Di Indonesia, keberhasilan proses pendidikan diukur dari kemampuan siswa dalam memahami dan mengaplikasikan materi pembelajaran. Terdapat berbagai metode penilaian yang dapat digunakan untuk mengukur kemampuan siswa, seperti soal pilihan ganda, jawaban singkat, dan esai (Salim dkk., 2022). Soal jawaban singkat mampu mengevaluasi kemampuan kognitif siswa, termasuk menghafal, merancang, dan mengungkapkan jawaban secara bebas berdasarkan pemikiran mereka (Hasanah dkk., 2018), lebih baik dibandingkan soal pilihan ganda. Namun, penilaian soal uraian ini memiliki tantangan tersendiri, terutama dalam hal konsistensi dan efisiensi (Burrows & D'Souza, 2005).

Sering kali guru ataupun dosen merasa kewalahan dengan banyaknya koreksi ulangan harian, ujian tengah semester, ataupun ujian akhir semester, terutama untuk guru yang memegang banyak kelas dan beban mengajar yang berlebih (Ardhana & Zahroh, 2023). Selain itu, adanya faktor subjektivitas pada proses penilaian manual dapat menyebabkan ketidakadilan dalam pemberian nilai. Ketidakakuratan ini tentunya dapat berdampak pada menurunnya motivasi siswa dan kepercayaan mereka terhadap sistem evaluasi. Oleh karena itu, diperlukan solusi yang dapat meningkatkan efisiensi sekaligus menjaga objektivitas penilaian soal jawaban singkat. Kemajuan teknologi dalam bidang kecerdasan buatan dan pemrosesan bahasa alami (*Natural Language Processing*) menawarkan peluang untuk mengotomatisasi proses penilaian soal jawaban singkat. Salah satu pendekatan yang saat ini sedang berkembang adalah *Automatic Short Answer Grading* (ASAG). ASAG merupakan proses penilaian jawaban singkat dalam bahasa alami terhadap pertanyaan objektif dengan menggunakan metode komputasional (Burrows dkk., 2015). Pendekatan ini menjadi solusi yang relevan untuk mendukung proses pendidikan di era digital.

Di Indonesia, pengembangan ASAG khususnya untuk soal uraian berbahasa Indonesia masih terbatas. Penelitian terdahulu menunjukkan bahwa sebagian besar pendekatan ASAG berfokus pada soal uraian berbahasa Inggris (Wilianto & Girsang, 2023; Ahmed dkk., 2022; Wangwiwattana & Tongvivat, 2022). Penelitian terdahulu juga hanya menyoroti hubungan antara dua masukan, yaitu referensi dan jawaban siswa (Wilianto & Girsang, 2023; Salim dkk., 2022; Wangwiwattana & Tongvivat, 2022), sehingga kurang optimal dalam menangkap hubungan semantik yang melibatkan soal sebagai konteks. Selain itu, permasalahan dalam ASAG semakin menantang ketika diterapkan pada lingkungan *open-domain*, di mana soal dan jawaban siswa tidak dibatasi pada topik tertentu. Dalam *open-domain*, model harus mampu menangani variasi linguistik yang jauh lebih luas, termasuk penggunaan sinonim, struktur kalimat kompleks, hingga istilah yang jarang digunakan. Oleh karena itu, dibutuhkan pendekatan berbasis *deep learning* yang tidak hanya memahami hubungan semantik antara referensi dan jawaban, tetapi juga mampu memanfaatkan konteks soal untuk menilai jawaban secara lebih akurat.

Teknologi pemrosesan bahasa alami berbasis transformer, seperti *Sentence BERT* (SBERT), menggunakan tiga masukan berupa soal, referensi, dan jawaban siswa menjadi solusi yang menjanjikan untuk mengatasi tantangan ini. Pemilihan SBERT sebagai representasi teks didasari oleh penelitian yang dilakukan oleh Condor dkk. (2021), yang mengeksplorasi proses *embedding* menggunakan SBERT, Word2Vec, dan *Bag of Words*. Hasilnya adalah SBERT secara konsisten mampu mengungguli kinerja Word2Vec dan *Bag of Words*. SBERT yang merupakan pengembangan dari model BERT (Reimers & Gurevych, 2019), mampu menghasilkan representasi kalimat yang lebih padat dan bermakna dibandingkan representasi token-level yang dihasilkan oleh BERT standar.

Meskipun model SBERT merupakan pendekatan berbasis *deep learning* yang mampu memberikan hasil yang memuaskan, sering kali diperlukan adaptasi lebih lanjut melalui teknik *fine-tuning*. *Fine-tuning* ini dilakukan untuk menyesuaikan model dengan tugas tertentu yang sedang dikerjakan. Salah satu metode yang dapat digunakan adalah *fine-tuning* menggunakan *triplet loss*. *Triplet loss* bekerja dengan mengoptimalkan jarak antara *embedding* dari pasangan data yang relevan (positif) dan yang tidak relevan (negatif) (Schroff dkk., 2015). Salah satu tantangan utama dalam menggunakan *triplet loss* ini adalah pemilihan sampel positif dan negatif dari jawaban siswa, mengingat data jawaban siswa merupakan data yang tidak berlabel (*unsupervised*).

Metode *clustering* digunakan untuk menjawab tantangan ini dengan mengelompokkan jawaban siswa berdasarkan tingkat kesamaan semantik. Metode ini menggunakan algoritma *K-Means Clustering* yang bekerja dengan menghitung *silhouette score* dari *embedding* jawaban siswa sehingga didapati nilai ' k ' (*cluster optimal*). *Cluster* yang paling dekat dengan referensi jawaban kemudian digunakan sebagai sampel positif, sedangkan *cluster* yang paling jauh dari referensi jawaban digunakan sebagai sampel negatif. Selain efisiensi dalam pengembangan, evaluasi sistem ASAG juga memegang peranan penting. Salah satu metrik yang relevan untuk tugas ini adalah SMAPE (*Symmetric Mean Absolute Percentage Error*), yang mengukur tingkat kesalahan prediksi dengan mempertimbangkan nilai absolut dan persentase (Kreinovich dkk., 2014). Penggunaan SMAPE memastikan bahwa model mampu memberikan prediksi yang akurat dan konsisten.

Dengan latar belakang tersebut, penelitian ini mengeksplorasi pendekatan ASAG pada soal uraian berbahasa Indonesia menggunakan SBERT dan metode *K-Means Clustering*. Penelitian ini dapat memberikan solusi praktis bagi dunia pendidikan, terutama dalam mengatasi tantangan penilaian soal uraian jawaban singkat yang selama ini menjadi hambatan. Dengan mengintegrasikan teknologi canggih berbasis AI, seperti transformer, penelitian ini juga bertujuan untuk meningkatkan efisiensi, akurasi, dan objektivitas dalam proses penilaian, serta mendukung upaya transformasi digital dalam sektor pendidikan di Indonesia.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan sebelumnya, rumusan masalah yang diangkat dalam penelitian ini adalah mengeksplorasi pendekatan ASAG pada soal uraian berbahasa Indonesia menggunakan SBERT dan metode *K-Means Clustering*. Adapun masalah khusus yang dibahas dalam penelitian ini adalah sebagai berikut.

1. Bagaimana pengaruh *fine-tuning* terhadap masing-masing *dataset* dalam model ASAG yang dikembangkan?
2. Bagaimana pembagian nilai α dan β dalam perhitungan *cosine similarity* terhadap masing-masing *dataset* dalam model ASAG yang dikembangkan?
3. Bagaimana pengaruh model SBERT yang di-*training* menggunakan korpus berbahasa Indonesia dan korpus multibahasa terhadap masing-masing *dataset* dalam model ASAG yang dikembangkan?

1.3 Tujuan dan Manfaat

Tujuan umum dari penelitian ini adalah menerapkan ASAG pada soal uraian berbahasa Indonesia menggunakan SBERT dan metode *K-Means Clustering*. Adapun tujuan khusus dari penelitian ini adalah sebagai berikut.

1. Mengetahui pengaruh *fine-tuning* terhadap masing-masing *dataset* dalam model ASAG yang dikembangkan.
2. Mencari nilai α dan β optimal dalam perhitungan *cosine similarity* terhadap masing-masing *dataset* dalam model ASAG yang dikembangkan.
3. Mengetahui pengaruh model SBERT yang di-*training* menggunakan korpus berbahasa Indonesia dan korpus multibahasa terhadap masing-masing *dataset* dalam model ASAG yang dikembangkan.

Manfaat dari penelitian ini adalah penilaian ujian uraian jawaban singkat berbahasa Indonesia dapat dilakukan secara otomatis tanpa memakan banyak waktu, serta pemberian nilai manual untuk pelatihan model dapat dilakukan seminimal mungkin. Diharapkan pula hasil penelitian ini dapat diterapkan secara luas untuk berbagai domain, serta mampu memberikan hasil penilaian yang lebih akurat, objektif, dan konsisten.

1.4 Ruang Lingkup

Ruang lingkup dalam penelitian *Automatic Short Answer Grading* pada soal uraian berbahasa Indonesia menggunakan SBERT dan metode *K-Means Clustering* ini dijabarkan sebagai berikut.

1. Penelitian menggunakan dua model SBERT yang telah di *fine-tuning*, yaitu model SBERT yang di-*training* menggunakan korpus berbahasa Indonesia (IndoSBERT) dan model yang di-*training* menggunakan korpus multibahasa (*paraphrase-multilingual-mpnet-base-v2*).
2. Nilai prediksi yang dihasilkan dari model ASAG berupa nilai bertipe *float* dengan rentang 0-1.
3. Penelitian ini tidak menjelaskan eksplorasi mengenai model *Pseudo-Siamese Neural Network* yang digunakan untuk assesmen jawaban.
4. *Tools* yang digunakan dalam penelitian ini adalah Google Colaboratory dengan menggunakan bahasa Python.

1.5 Sistematika Penulisan

Berikut adalah sistematika penulisan laporan skripsi *Automatic Short Answer Grading* pada soal uraian berbahasa Indonesia menggunakan SBERT dan metode *K-Means Clustering*, untuk memudahkan pembaca dalam memahami laporan secara urut dan jelas.

BAB I PENDAHULUAN

Bab ini menyajikan aspek penting yang menjadi dasar dalam penyusunan laporan skripsi. Adapun hal-hal yang dibahas meliputi latar belakang, rumusan masalah, tujuan dan manfaat, ruang lingkup, serta sistematika penulisan.

BAB II LANDASAN TEORI

Bab ini memuat tinjauan pustaka dan teori-teori yang digunakan dalam penelitian. Adapun hal-hal yang dibahas meliputi *state of the art*, *Automatic Short Answer Grading*, *regular expressions*, pra-pemrosesan data yang mencakup *data formatting*, normalisasi data, pembersihan data, serta *case folding*, *K-Means Clustering*, *silhouette score*, *fine-tuning* menggunakan *triplet loss*, *forward* dan *backward pass*, *sentence embedding*, *mean pooling*, model *pre-trained* yang mencakup BERT, IndoSBERT, XLM-RoBERTa, dan *paraphrase-multilingual-mpnet-v2*, perhitungan *cosine similarity*, *few-shot learning*, dan evaluasi SMAPE.

BAB III METODOLOGI PENELITIAN

Bab ini menyajikan alur serta tahapan-tahapan yang digunakan sebagai acuan dalam melaksanakan penelitian. Adapun hal-hal yang dibahas meliputi gambaran umum penelitian, pengumpulan data, pra-pemrosesan data, *clustering*, *fine-tuning*, *sentence embedding*, perhitungan *cosine similarity*, pembagian data, pembangunan model assesmen, evaluasi model menggunakan SMAPE, serta *experimental setup*.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil dan analisis dari skenario serta eksperimen yang dilakukan dalam penelitian sesuai dengan metodologi penelitian yang dijadikan sebagai acuan.

BAB V PENUTUP

Bab ini menyajikan kesimpulan yang didapat dari hasil eksperimen yang telah dilakukan dalam penelitian serta saran untuk penelitian selanjutnya.