

BAB I

PENDAHULUAN

Bab pendahuluan ini membahas mengenai latar belakang masalah, rumusan masalah, tujuan dan manfaat, ruang lingkup, serta sistematika penulisan yang akan digunakan dalam dokumen skripsi ini.

1.1 Latar Belakang

Salah satu elemen krusial dalam proses pembelajaran adalah *assessment* atau evaluasi (Prabhudesai & Duong, 2019). Sistem evaluasi digunakan untuk mengukur keterampilan berpikir siswa, khususnya dalam menjawab pertanyaan melalui proses penilaian atas jawaban yang diberikan (Susanti dkk., 2022). Evaluasi penting untuk dilakukan karena bertujuan untuk mengetahui sejauh mana pengetahuan telah berhasil diserap oleh siswa dan sejauh mana mereka memahami topik yang diajarkan. Evaluasi ini dapat dilakukan dengan berbagai jenis tes, seperti pilihan ganda, uraian singkat, dan esai.

Berbeda dengan pilihan ganda yang memiliki kunci jawaban tunggal dan mudah dinilai, penilaian terhadap soal uraian singkat dan esai lebih kompleks dan membutuhkan evaluasi yang lebih mendalam. Pada jenis soal ini, setiap siswa dapat memberikan jawaban dengan cara yang berbeda, tergantung pada gaya ekspresi mereka, sehingga menambah tantangan bagi pengajar untuk melakukan penilaian secara konsisten dan objektif (Chang dkk., 2024). Evaluasi manual ini menjadi tugas yang tidak hanya memakan banyak waktu, tetapi juga sering kali terasa monoton dan membosankan, mengingat banyaknya jumlah siswa yang harus dinilai (Wilianto & Girsang, 2023).

Untuk mengatasi tantangan tersebut, pendekatan *Automatic Short Answer Grading* (ASAG) dapat menjadi solusi yang efektif. Sistem ASAG meringankan beban pengajar dengan mengotomasi proses penilaian jawaban singkat, sekaligus mengurangi potensi bias manusia yang sering muncul dalam penilaian manual (Prabhudesai & Duong, 2019). Berbeda dengan metode tradisional, di mana penilaian dapat dipengaruhi oleh persepsi dan interpretasi pribadi, ASAG menggunakan algoritma untuk mengevaluasi kesamaan semantik antara jawaban siswa dan jawaban yang benar (Tulu dkk., 2021). Pendekatan ini

memungkinkan penilaian yang lebih konsisten dan objektif, karena didasarkan pada pemahaman makna atau konteks jawaban, bukan hanya kata-kata yang digunakan.

Untuk menghasilkan prediksi akhir dalam tugas ASAG, berbagai pendekatan telah dirancang untuk menangkap kesesuaian antara jawaban siswa dan jawaban referensi. Salah satu metode yang sering digunakan adalah *cosine similarity* (Bonthu dkk., 2023; Wilianto & Girsang, 2023), yang menghitung tingkat kemiripan antara dua representasi vektor berdasarkan sudut di antara mereka. Pendekatan berbasis *linear regression* (Haidir & Purwarianti, 2020) juga dapat digunakan untuk memetakan hubungan antara fitur masukan dan skor evaluasi. Selain itu, pendekatan berbasis *neural network* (Chang dkk., 2024; Prabhudesai & Duong, 2019) telah menjadi solusi yang semakin populer dalam ASAG. Salah satu implementasi yang umum digunakan adalah *Bidirectional Long Short-Term Memory* (BiLSTM) (Gomaa dkk., 2023; Patil & Adhiya, 2023). BiLSTM mampu menangkap konteks dari kedua arah dalam data sekuensial, sehingga meningkatkan akurasi dalam penilaian jawaban siswa.

Namun, salah satu tantangan utama dalam implementasi ASAG di dunia nyata adalah keterbatasan jumlah data yang tersedia. Kondisi ini membuat pengembangan model ASAG yang mampu memberikan performa baik meski dengan data latih yang terbatas menjadi tantangan besar. Untuk mengatasi masalah ini, pendekatan *few-shot learning* telah mulai diterapkan dalam pengembangan sistem ASAG. Pendekatan ini memungkinkan model untuk belajar dari sejumlah kecil data berlabel dengan cara mengoptimalkan kemampuan generalisasi model. Meskipun pendekatan ini menunjukkan potensi besar, penelitian terkait *few-shot learning* untuk jawaban singkat berbahasa Indonesia masih tergolong langka, membuka peluang luas untuk eksplorasi lebih lanjut di bidang ini.

Few-shot learning dengan pendekatan *transfer learning* telah menunjukkan efektivitasnya pada skenario dengan keterbatasan *dataset* (Chang dkk., 2024). Strategi yang digunakan adalah melatih model *external knowledge* pada *dataset* besar, seperti kumpulan teks Wikipedia. Model yang telah dilatih ini kemudian diadaptasi untuk tugas spesifik melalui *fine-tuning* pada *dataset* yang lebih kecil, sehingga memungkinkan model untuk menggeneralisasi dengan lebih baik meskipun hanya tersedia sedikit data. Meskipun metode ini terbukti efektif, penelitian ini masih menggunakan data latih sebesar 70%, yang tidak

sepenuhnya mencerminkan tantangan dunia nyata di mana data latih yang tersedia jauh lebih terbatas.

Oleh karena itu, penelitian ini mengadopsi metode tersebut untuk mengembangkan model ASAG berbasis *few-shot learning* dengan memanfaatkan pendekatan *transfer learning*, dengan penekanan khusus pada penggunaan proporsi data latih yang sangat kecil. Pendekatan ini dipilih sebagai solusi untuk mengatasi tantangan keterbatasan data serta kebutuhan akan model ASAG yang *open-domain*. Untuk mendukung proses *transfer learning*, model berbasis pengetahuan eksternal akan dilatih terlebih dahulu menggunakan *dataset* Wikipedia Bahasa Indonesia. *Dataset* ini dipilih karena cakupan domainnya yang luas dan mencakup berbagai topik yang relevan untuk tugas ASAG dalam konteks *open-domain*, memungkinkan model untuk menangkap pengetahuan semantik yang lebih kaya dan kontekstual, serta meningkatkan kemampuan model untuk menggeneralisasi meskipun data latih terbatas.

Selain itu, penelitian ini mengadopsi model *Pseudo-Siamese Neural Network* (PSNN) yang telah dikembangkan sebelumnya dengan modifikasi yang disesuaikan untuk tugas ASAG pada penelitian ini. PSNN dipilih karena struktur dua kepala (*head*) nonidentik pada arsitekturnya (Chang dkk., 2024), memungkinkan model untuk mengintegrasikan informasi antara jawaban siswa dan jawaban referensi serta jawaban siswa dan soal. Model ini dimodifikasi dengan *weighted fusion* untuk lebih menitikberatkan penilaian pada kesamaan jawaban siswa dan jawaban referensi.

Sementara itu, model *Siamese BiLSTM* (Prabhudesai & Duong, 2019) dan *BiLSTM-Parallel CNN* (Kaya & Cicekli, 2024) digunakan sebagai pembanding untuk PSNN. *Siamese BiLSTM* dipilih karena kemampuannya dalam menangkap hubungan sekuensial dalam teks dan untuk melihat pengaruh jaringan *siamese* dalam menghasilkan prediksi. Adapun model *BiLSTM-Parallel CNN* dipilih karena menggabungkan keunggulan *BiLSTM* dalam memahami hubungan sekuensial dengan kekuatan *CNN* paralel dalam menangkap fitur lokal dalam berbagai skala.

Berdasarkan tinjauan literatur, ketiga model yang disebutkan di atas dipilih sebagai pendekatan regresi berbasis *neural network* yang sesuai untuk tugas ASAG yang dikembangkan. Model-model ini telah terbukti memberikan hasil yang baik dalam penelitian sebelumnya, sehingga diharapkan dapat menghasilkan performa yang optimal dalam

skenario yang diuji. Pemilihan ketiga model dengan arsitektur yang berbeda ini juga didasarkan pada tujuan penelitian untuk mengidentifikasi model yang paling efektif diterapkan di dunia nyata, terutama dalam kondisi keterbatasan data. Dengan demikian, penelitian ini dapat memberikan rekomendasi yang tepat bagi implementasi ASAG dan diharapkan dapat memberikan wawasan mendalam mengenai efektivitas masing-masing pendekatan dalam konteks pengembangan ASAG yang akurat.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan sebelumnya, rumusan masalah yang diangkat dalam penelitian ini adalah bagaimana pengembangan ASAG berbasis *few-shot learning* dengan menggunakan *Pseudo-Siamese Neural Network*, *Siamese BiLSTM Neural Network*, dan *BiLSTM-Parallel CNN*. Adapun masalah khusus yang akan dibahas dalam penelitian ini meliputi:

1. Bagaimana pengaruh variasi proporsi persentase data latih dan data uji terhadap performa ketiga model tersebut?
2. Bagaimana perbandingan performa antara *Pseudo-Siamese Neural Network*, *Siamese BiLSTM Neural Network*, dan *BiLSTM-Parallel CNN* dalam tugas ASAG?

1.3 Tujuan dan Manfaat

Tujuan umum penelitian ini adalah menerapkan *few-shot learning* untuk tugas ASAG pada uraian jawaban singkat berbahasa Indonesia. Tujuan khusus penelitian ini antara lain:

1. Menganalisis pengaruh proporsi persentase data latih dan data uji terhadap performa ketiga model untuk memahami efektivitas pendekatan *few-shot learning* dalam kondisi *dataset* yang terbatas.
2. Membandingkan performa *Pseudo-Siamese Neural Network*, *Siamese BiLSTM Neural Network*, dan *BiLSTM-Parallel CNN* dalam mengimplementasikan ASAG berbasis *few-shot learning*.

Manfaat penelitian ini diharapkan dapat memberikan kontribusi nyata dalam berbagai aspek kehidupan, khususnya dalam dunia pendidikan. Sistem ASAG berbasis *few-shot learning* yang dihasilkan dari penelitian ini dapat mendukung efisiensi proses penilaian jawaban siswa, terutama untuk tugas dengan jumlah siswa yang besar, sehingga dapat mengurangi beban kerja guru atau dosen dalam menilai secara manual. Dengan fokus pada

pengembangan sistem untuk jawaban singkat berbahasa Indonesia, penelitian ini juga menjadi langkah penting dalam memperluas implementasi teknologi berbasis kecerdasan buatan di bidang pendidikan dalam konteks lokal, yang selama ini kurang tereksplorasi dibandingkan bahasa global lainnya.

1.4 Ruang Lingkup

Ruang lingkup dalam menerapkan *few-shot learning* untuk tugas ASAG pada penelitian ini adalah sebagai berikut:

1. Pada penelitian ini tidak dilakukan eksplorasi mengenai representasi semantik.
2. Keluaran dari model ASAG berupa prediksi nilai bertipe *float* dengan rentang 0-1.
3. *Tools* yang digunakan dalam penelitian ini adalah Google Colaboratory dengan menggunakan bahasa Python.

1.5 Sistematika Penulisan

Sistematika penulisan digunakan untuk memberikan gambaran yang runtut dan jelas mengenai penyusunan laporan skripsi yang berjudul Pendekatan *Few-Shot Learning* untuk *Automatic Short Answer Grading*: Implementasi *Pseudo-Siamese Neural Network*, *Siamese BiLSTM*, dan *BiLSTM-Parallel CNN*. Sistematika penulisan laporan akan dijelaskan sebagai berikut:

BAB I PENDAHULUAN

Bab ini membahas mengenai beberapa aspek penting, termasuk latar belakang, rumusan masalah, tujuan dan manfaat, ruang lingkup, dan sistematika penulisan laporan yang dibuat.

BAB II LANDASAN TEORI

Bab ini membahas mengenai *state-of-the-art* dan teori-teori yang menjadi landasan penelitian ini, meliputi pemrosesan bahasa alami, *Automatic Short Answer Grading* (ASAG), *regular expression*, prapemrosesan data tabular, prapemrosesan teks, *sentence embedding*, model IndoSBERT, *cosine similarity*, *siamese network*, *pseudo-siamese network*, *Convolutional Neural Network* (CNN), *Long Short-Term Memory* (LSTM), *Bidirectional LSTM* (BiLSTM), *Few-shot Learning* (FSL), *transfer learning*, *fine-tuning*, dan *Symmetric Mean Absolute Percentage Error* (SMAPE).

BAB III METODOLOGI PENELITIAN

Bab ini membahas mengenai metodologi penelitian skripsi, meliputi pengumpulan data, prapemrosesan data, *sentence embedding*, pembagian data, pengembangan model ASAG, dan pengujian model ASAG.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil penelitian serta analisis terkait performa model yang dikembangkan. Fokus utama dalam bab ini adalah hasil pengolahan data, evaluasi performa model *Pseudo-Siamese Neural Network*, *Siamese BiLSTM*, dan *BiLSTM-Parallel CNN*, serta perbandingan kinerja ketiga model dalam dua skenario pembagian data latih sekaligus validasi, yaitu 20% dan 10%.

BAB V KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan dari seluruh bab yang telah dibahas dan saran sebagai bahan masukan untuk pengembangan atau penelitian selanjutnya.