

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Lingkungan pembelajaran cerdas atau *Smart Learning Environment* (SLE) didefinisikan sebagai sistem pembelajaran *hybrid* (mode *online*, *offline*, maupun campuran) yang memberikan proses belajar yang menyenangkan kepada siswa sambil mencapai hasil belajar dengan penggunaan alat dan teknik cerdas [1]. Tujuan utama dari pengembangan SLE adalah agar proses pembelajaran menjadi lebih efektif terutama dalam hal penggunaan sumberdaya dan waktu. Salah satu dari enam komponen utama dalam SLE adalah modul pembelajar (*Learner Module*) yang berisi catatan-catatan hasil penilaian atau evaluasi dari hasil pembelajaran siswa. Dalam kaitannya dengan hasil penilaian terhadap hasil pembelajaran siswa, SLE hadir untuk membantu menciptakan penilaian yang lebih berkeadilan bagi setiap siswa.

Penilaian merupakan suatu tahapan penting dalam proses pembelajaran di dunia pendidikan untuk mengukur sejauh mana pemahaman seorang siswa terhadap materi yang diberikan [2]. Penilaian dapat dilakukan dengan memberikan tes atau ujian kepada siswa. Pada umumnya terdapat dua jenis ujian yang dapat digunakan untuk melakukan penilaian, yaitu ujian pilihan ganda dan ujian esai.

Ujian esai merupakan salah satu jenis proses evaluasi pembelajaran yang diberikan dalam bentuk ujian tertulis, yang susunannya terdiri atas pertanyaan-pertanyaan yang mengandung permasalahan dan menuntut jawaban siswa melalui uraian kata yang merefleksikan kemampuan berpikir siswa [3]. Proses penilaian ujian esai tidaklah mudah dikarenakan jawabannya sangat bervariasi. Semakin banyak jumlah soal ujian dan siswa yang mengikuti ujian, maka semakin banyak pula jawaban yang harus dikoreksi oleh guru. Hal ini akan menyebabkan pekerjaan guru menjadi tidak efektif dan efisien karena waktu yang diperlukan relatif lama dan memerlukan konsentrasi serta tenaga lebih, sehingga dapat memungkinkan terjadinya *human error* jika masih dilakukan secara manual [4].

Penelitian terkait penilaian teks esai secara otomatis umumnya terbagi menjadi dua subjek utama, yaitu *Automated Essay Scoring* (AES) dan *Automated Short-*

Answer Scoring (ASAS) [5], [6]. Pengertian esai yang dimaksud dalam AES adalah karangan prosa tentang topik tertentu yang umumnya bersifat analitik, spekulatif, atau interpretatif [7], sedangkan pengertian jawaban singkat dalam ASAS adalah jawaban yang disusun dalam satu kalimat atau lebih dan umumnya sampai satu atau dua paragraf [8]. Penelitian ini menggunakan istilah jawaban esai untuk mengacu pada esai yang dimaksud dalam ASAS.

ASAS lebih menekankan penilaian jawaban pada aspek konten jawaban. Siswa dapat menjawab dalam berbagai bentuk susunan kata atau kalimat yang berbeda secara sintaksis maupun leksikal, tetapi memiliki makna yang sama dengan kunci jawaban [9]. Jawaban siswa memungkinkan menggunakan sinonim, kata-kata dan pernyataan polisemi yang merupakan parafrasa satu sama lain dari kunci jawaban yang telah ditetapkan. Semakin tinggi tingkat kemiripan makna atau semantik jawaban siswa dengan kunci jawaban, maka akan semakin tinggi skor penilaian yang diperoleh oleh siswa tersebut.

Penilaian esai otomatis pada dasarnya mengukur kemiripan antara dua teks. Pengukuran kesamaan teks dapat dilakukan dengan dua jenis pendekatan, yaitu kesamaan leksikal dan kesamaan semantik [10]. Suatu *string* memiliki kesamaan leksikal yang mirip jika memiliki karakter atau urutan kata yang mirip. Jarak antara keduanya dihitung berdasarkan karakter atau urutan kata [11]. *String* juga memiliki kesamaan semantik yang mirip jika memiliki makna yang sama, tidak berlawanan satu sama lain, digunakan dengan cara yang sama, dan dalam konteks yang sama. Jarak antara keduanya dihitung berdasarkan kemiripan maknanya [12].

Kesamaan leksikal dapat dihitung dengan menggunakan ukuran berbasis *string* dengan memanfaatkan algoritma seperti *cosine similarity*, *jaccard coefficient*, dan *longest common subsequence*. *Cosine similarity* merupakan ukuran kemiripan antara dua vektor (vektor dokumen dan vektor *query*) yang dibandingkan dalam suatu ruang dimensi yang diperoleh dari nilai kosinus sudut dari perkalian dua vektor tersebut [13]. *Jaccard coefficient* adalah suatu algoritma yang berfungsi untuk membandingkan dua dokumen berdasarkan kata yang dimilikinya [14]. *Longest common subsequence* adalah metrik kesamaan berbasis karakter yang menghitung kesamaan antara dua *string* berdasarkan panjang karakter terpanjang (*subsequence*) yang ada di kedua *string* tersebut [10]. Sementara kesamaan

semantik dapat dihitung menggunakan algoritma *knowledge-based similarity*, *corpus-based similarity*, dan *word embedding*.

Knowledge-based similarity adalah suatu algoritma yang mengambil informasi dari jaringan semantik yang merepresentasikan hubungan antar kata untuk mengukur tingkat kemiripan suatu teks. Jaringan semantik yang banyak digunakan untuk mendeteksi kemiripan kata adalah *WordNet*. *WordNet* merupakan basis data leksikal berbahasa Inggris yang menghimpun relasi semantik yang dimiliki oleh suatu kata seperti sinonim, antonim, dan hiponim [15]. Penggunaan *WordNet* mampu menggambarkan hubungan semantik suatu kata secara lengkap dan dapat dimanfaatkan untuk mengukur kemiripan semantik antar kata atau kalimat. Akan tetapi, basis data *WordNet* belum tersedia secara lengkap untuk beberapa bahasa, misalnya bahasa Indonesia, sehingga kurang cocok untuk diterapkan dalam mengukur kemiripan teks berbahasa Indonesia.

Corpus-based similarity digunakan untuk menghitung kemiripan teks dengan mempertimbangkan kemiripan makna melalui kamus. Terdapat beberapa jenis algoritma *Corpus-based similarity*, salah satunya yaitu *Latent Semantic Analysis* (LSA). LSA dibentuk melalui matriks *term-document* suatu korpus teks. LSA tidak memerlukan sumber daya eksternal untuk menghitung kesamaan semantik seperti yang dilakukan oleh algoritma *Knowledge-based similarity*. LSA dalam sistem penilaian esai otomatis digunakan untuk mengekstraksi topik yang tersirat pada setiap jawaban siswa dan referensi kunci jawaban. Perhitungan kemiripan semantik antara jawaban siswa dan referensi jawaban dihitung menggunakan *cosine similarity* [16].

Algoritma selanjutnya yang dapat digunakan untuk mengukur kesamaan semantik adalah algoritma penyematan kata (*word embedding*). Algoritma ini merepresentasikan suatu kata ke dalam sebuah vektor yang berisi informasi semantik dari suatu kata. Terdapat beberapa jenis model *word embedding*, salah satunya adalah *Bidirectional Encoder Representations from Transformers* (BERT). Algoritma BERT adalah suatu model bahasa *pre-trained* yang dirilis oleh tim *Google AI Language* pada tahun 2018, yang dimana telah mencapai hasil yang sesuai dalam berbagai kasus NLP [17]. Dengan BERT, model bahasa dapat

mempelajari konteks kata berdasarkan kata-kata yang mengelilinginya, tidak hanya kata yang mendahului atau mengikutinya.

Penelitian sebelumnya yang dilakukan oleh Hasanah *et al.* [18] telah berhasil mengembangkan penilaian jawaban esai otomatis dengan pendekatan teks berbasis *string* dan pencocokan kata kunci tanpa mempertimbangkan semantik kata. Akan tetapi, untuk setiap pertanyaan yang ditanyakan untuk mengantisipasi keanekaragaman jawaban siswa diperlukan beberapa kunci jawaban. Kemudian, penelitian selanjutnya yang dilakukan oleh Lubis *et al.* [19] mencoba mengatasi kekurangan dari penelitian Hasanah *et al.* [18] yaitu dengan membuat model penilaian jawaban esai otomatis berdasarkan kesamaan semantik berbasis *word embedding* menggunakan *Word2Vec* dan juga menambahkan analisis sintaksis untuk menentukan makna kalimat yang berlawanan. Penelitian tersebut cukup menggunakan satu kunci jawaban saja untuk dapat melakukan penilaian jawaban esai otomatis. Akan tetapi, belum mampu memberikan penilaian secara tepat seperti manusia saat mencocokkan jawaban yang membutuhkan akurasi urutan kata.

Berdasarkan latar belakang yang telah disampaikan, maka penelitian ini akan melanjutkan penelitian yang dilakukan oleh Lubis *et al.* [19] dengan tujuan untuk meningkatkan akurasi penilaian jawaban esai otomatis menggunakan kesamaan semantik dan analisis sintaksis. Penggantian model *word embedding* dari yang sebelumnya menggunakan *Word2Vec* menjadi BERT berbahasa Indonesia, yaitu IndoBERT dan penambahan analisis sintaksis diharapkan dapat meningkatkan hasil akurasi penilaian jawaban esai secara otomatis. Perhitungan kesamaan semantik antara jawaban siswa dan referensi kunci jawaban guru dilakukan dengan cara merepresentasikan jawaban ke dalam vektor dokumen dengan memanfaatkan algoritma IndoBERT. Model ini dapat digunakan untuk menghitung rata-rata vektor kata yang menyusun dokumen. Perhitungan *cosine similarity* digunakan untuk menghitung jarak antar vektor dokumen dari setiap pasangan jawaban siswa dan kunci jawaban [9]. Nilai kesamaan semantik antara jawaban siswa dan kunci jawaban guru direpresentasikan oleh jarak antar vektor dokumen tersebut. Selain itu, akan ditambahkan analisis sintaksis juga yang bertujuan untuk membandingkan kebenaran antar kalimat dan mendeteksi soal yang masih membutuhkan urutan kata

pada kalimat jawaban sebagai pertimbangan penilaian dengan memanfaatkan *part-of-speech tagging* dan *dependency parsing*.

1.2 Rumusan Masalah

Berdasarkan latar belakang permasalahan di atas, maka dapat dirumuskan permasalahan pada penelitian ini adalah sebagai berikut :

1. Bagaimana mengimplementasikan algoritma BERT berbahasa Indonesia untuk pengukuran kesamaan semantik dalam penilaian esai otomatis?
2. Bagaimana menambahkan analisis sintaksis untuk mendeteksi urutan kata pada kalimat jawaban?
3. Bagaimana hasil nilai MAE dari sistem penilaian esai otomatis menggunakan algoritma BERT berbahasa Indonesia untuk pengukuran kesamaan semantik dan analisis sintaksis untuk mendeteksi urutan kata pada kalimat jawaban?

1.3 Tujuan Penelitian

Tujuan yang hendak dicapai dalam penelitian ini adalah sebagai berikut :

1. Mengimplementasikan algoritma BERT berbahasa Indonesia untuk pengukuran kesamaan semantik dalam penilaian esai otomatis?
2. Menambahkan analisis sintaksis untuk mendeteksi urutan kata pada kalimat jawaban
3. Mengetahui hasil nilai MAE dari sistem penilaian esai otomatis menggunakan algoritma BERT berbahasa Indonesia untuk pengukuran kesamaan semantik dan analisis sintaksis untuk mendeteksi urutan kata pada kalimat jawaban

1.4 Manfaat Penelitian

Adapun manfaat yang didapatkan dari sistem ini, diantaranya :

1. Membantu guru atau pengajar dalam melakukan penilaian terhadap soal esai secara otomatis sehingga dapat meningkatkan objektivitas dan efisiensi waktu
2. Menjadi referensi penelitian selanjutnya mengenai pengembangan atau penggunaan algoritma BERT berbahasa Indonesia untuk pengukuran kesamaan semantik dan analisis sintaksis dalam melakukan penilaian esai otomatis

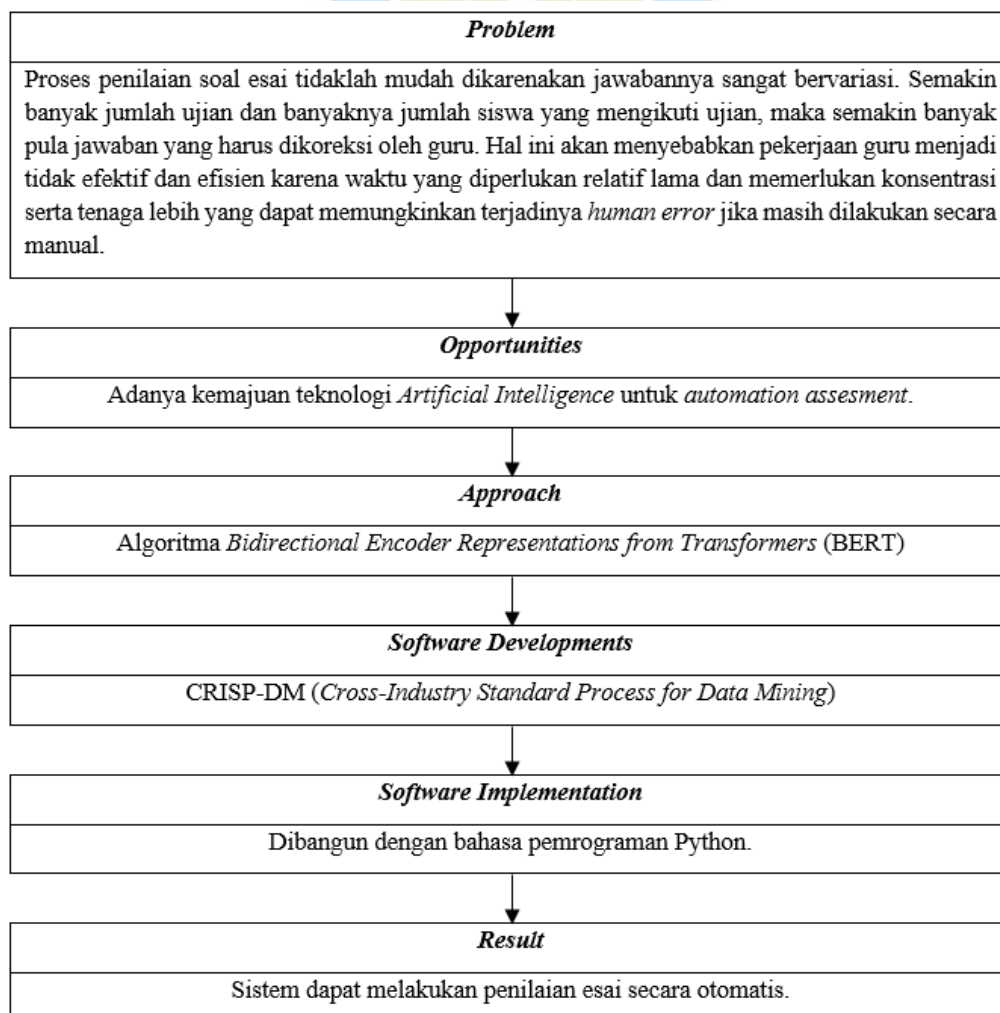
1.5 Batasan Masalah

Pada penelitian ini ditetapkan beberapa batasan masalah diantaranya :

1. Penelitian ini menggunakan *dataset* dari soal, kunci jawaban, dan jawaban esai berbahasa Indonesia
2. Penelitian ini menggunakan *dataset* dari soal, kunci jawaban, dan jawaban esai mata pelajaran pemrograman dasar kelas 10
3. Penelitian ini hanya menguji jawaban teks, tidak berupa file gambar, suara, perhitungan matematika ataupun karakter
4. Menggunakan *library* sastrawi yang menerapkan algoritma Nazief & Adriani untuk melakukan proses *stemming*
5. Menggunakan *library* spacy stanza untuk melakukan *part-of-speech tagging* dan *dependency parsing*

1.6 Kerangka Pemikiran

Kerangka pemikiran dari penelitian ini dapat diuraikan pada Gambar 1.1.



Gambar 1. 1 Kerangka Pemikiran

1.7 Sistematika Penulisan

Dalam penulisan ini dilakukan pembahasan yang terfokus pada inti permasalahan. Untuk itu penulis menyusun sistematika penulisan tugas akhir ini sebagai berikut :

BAB I : PENDAHULUAN

Pada bab ini penulis akan menjelaskan tentang latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, kerangka pemikiran, dan sistematika penulisan.

BAB II : KAJIAN LITERATUR

Pada bab ini menerangkan tentang hasil kajian literatur yang akan mendasari penelitian. Secara garis besar, bab ini menjelaskan mengenai penelitian-penelitian sebelumnya terkait penilaian esai otomatis dan berbagai konsep yang berkaitan dengan topik penelitian.

BAB III : METODOLOGI PENELITIAN

Pada bab ini penulis akan membahas mengenai metode yang diterapkan pada penelitian. Pada bab ini juga dirumuskan tahap-tahap apa saja yang akan dilaksanakan selama penelitian.

BAB IV : HASIL DAN PEMBAHASAN

Pada bab ini penulis akan membahas mengenai hasil dan temuan berdasarkan langkah-langkah pada metodologi penelitian. Hasil ini akan membahas mengenai bagaimana penilaian esai otomatis menggunakan algoritma IndoBERT untuk pengukuran kesamaan semantik dan analisis sintaksis untuk mendeteksi urutan kata pada kalimat jawaban akan bekerja.

BAB V : PENUTUP

Pada bab ini akan menarik kesimpulan dari penelitian yang sudah dilakukan dan akan ditampilkan penyampaian saran untuk penelitian terkait kedepannya.