

BAB 1

PENDAHULUAN

1.1. Latar Belakang

Evaluasi terhadap proses pendidikan formal diperlukan untuk mengetahui capaian keahlian dan tingkat pemahaman peserta didik mengenai suatu disiplin ilmu. Metode yang digunakan untuk mengevaluasi pemahaman peserta didik dapat melalui ujian. Ujian ada beberapa macam bentuk soalnya. Salah satu bentuk soal yang sering digunakan yaitu pilihan ganda (*multiple choice*) dan isian (*essay*) [1].

Penilaian ujian berbentuk soal pilihan ganda mudah dilakukan, cukup dengan pencocokan jawaban mahasiswa dan jawaban kunci. Namun pada penilaian soal berbentuk *essay* memerlukan proses yang relatif lama. Sebagai contoh, rata-rata panjang satu jawaban *essay* pada mata kuliah Organisasi Komputer di Teknik Informatika sebanyak 30 kata. Ketika ujian tengah semester digunakan 10 soal *essay* sehingga terdapat sekitar 300 kata untuk satu mahasiswa. Jumlah mahasiswa dalam satu angkatan tahun 2024 yang mengambil mata kuliah tersebut sebanyak 96 mahasiswa. Sehingga rata-rata total kata sebanyak 28.800 kata. Kecepatan membaca efektif untuk tingkat perguruan tinggi berada pada rentang 300-350 kata per menit [2]. Jadi waktu yang diperlukan untuk mengoreksi jawaban *essay* selama 1 jam 42 menit. Pendidik juga perlu memeriksa satu per satu lembar jawab soal isian kemudian dicocokkan dengan jawaban yang ditentukan oleh pendidik tersebut. Selain memerlukan proses yang lama, kelemahan proses penilaian dengan bentuk soal *essay* rawan terhadap hasil yang tidak ajek seperti perbedaan nilai meskipun jawaban memiliki makna yang mirip karena faktor subjektivitas, salah satu penyebab yaitu jawaban mahasiswa yang bervariasi terhadap soal dan terdapat faktor sinonim dan kesamaan semantik dari kata terhadap jawaban *essay* yang sesungguhnya [3].

Berangkat dari masalah tersebut perlunya sebuah mekanisme asesmen ujian berbasis soal *essay*. Teknologi informasi dapat diimplementasikan terhadap masalah tersebut dengan metode pengolahan data teks dan mengukur tingkat kemiripan jawaban mahasiswa dengan jawaban kunci lalu kemudian melakukan

penilaian (*scoring*). Pengolahan data jawaban essay akan memerlukan pemrosesan bahasa alami sebab bentuk data yang akan diolah merupakan data teks sehingga salah satu metode yang dapat digunakan adalah sub bidang kecerdasan buatan untuk mengolah bahasa yaitu *Natural Language Processing* (NLP). Model NLP saat ini semakin berkembang setelah arsitektur *transformer* dicetuskan pada tahun 2017 dan memiliki kemampuan memahami konteks kata dalam sebuah kalimat utuh melalui mekanisme *attention* [4].

Penelitian ini perlu dilakukan sebab dapat memberikan kontribusi terhadap dunia pendidikan terutama pendidikan tinggi terkait efisiensi proses asesmen ujian berbasis *essay*. Soal esai tetap perlu digunakan dalam pendidikan sebab pendidik (dosen) dapat menilai jawaban mahasiswa secara komprehensif dan tidak hanya dengan jawaban benar-salah seperti bentuk soal pilihan ganda (*multiple choice*). Dengan demikian maka dosen dapat mengetahui pemahaman mahasiswa terhadap materi kuliah yang telah diberikan melalui hasil ujian berbasis esai yang telah ditingkatkan efisiensi proses asesmennya.

1.2. Perumusan Masalah

Masalah yang dirumuskan berdasarkan latar belakang yang dipaparkan adalah sebagai berikut.

- a. Bagaimana menerapkan model NLP pada penilaian jawaban esai Bahasa Indonesia?
- b. Bagaimana performa model NLP untuk melakukan tugas penilaian jawaban *essay* menggunakan pengukuran *Quadratic Weighted Kappa* dan *Root Mean Square Error*?

1.3. Tujuan Penelitian

Tujuan dilaksanakannya penelitian adalah sebagai berikut.

- a. Menerapkan model NLP pada penilaian jawaban *essay* Bahasa Indonesia.
- b. Mengetahui performa model NLP untuk melakukan tugas penilaian jawaban *essay*.

1.4. Batasan Masalah

Ruang lingkup dalam penelitian perlu dibatasi agar pembahasan dapat lebih terperinci. Batasan masalah dalam penelitian adalah sebagai berikut.

- a. Jawaban esai menggunakan bahasa Indonesia dan berbentuk teks digital tanpa mengandung angka untuk menanggulangi kosakata yang tidak dikenali (*out of vocabulary/OOV*) oleh model NLP ketika proses tokenisasi kata.
- b. Panjang jawaban essay yang diproses dengan NLP tidak lebih dari 200 kata. Angka tersebut diperoleh karena model IndoBERT dapat memroses teks input dengan maksimal 512 token. Model IndoBERT menggunakan tokenizer yang memecah kata berimbuhan dan diasumsikan setiap kata dalam Bahasa Indonesia memiliki imbuhan di awal dan/atau akhir, sehingga akan menambah jumlah token yang diproses. Untuk menanggulangi jumlah token yang melebihi batas maksimum yang dapat diproses IndoBERT, maka perlu dilakukan pembatasan jumlah kata dalam teks input.
- c. Jawaban essay merupakan jawaban yang berbasis materi mata kuliah, bukan merupakan jawaban yang bersifat argumentatif atau hasil berpikir kritis.
- d. Topik yang digunakan sebagai simulasi soal dan jawaban essay merupakan tema pada ruang lingkup keilmuan Teknik Informatika.

1.5. Manfaat Penelitian

Manfaat dari penelitian yang dilakukan adalah sebagai berikut.

- a. Memberikan kontribusi terhadap dunia pendidikan mengenai efisiensi penilaian jawaban essay sehingga memudahkan pekerjaan pendidik (dosen) untuk menilai capaian peserta didik (mahasiswa) secara komprehensif melalui ujian essay.
- b. Metode NLP dapat digunakan untuk mengolah bahasa alami dalam hal ini Bahasa Indonesia dengan demikian program komputer dapat

memroses jawaban esai bukan hanya secara frekuensi kata kunci yang terkandung dalam jawaban essay tersebut melainkan juga mampu menilai secara semantik (makna kalimat).

