

BAB I

PENDAHULUAN

1.1 Latar Belakang

Ujian merupakan salah satu alat penilaian dalam proses belajar mengajar yang berfungsi memberikan penilaian menyeluruh terhadap mutu pembelajaran dan mengukur kinerja siswa. Di sisi lain, ujian menjadi salah satu beban tambahan dan tanggung jawab yang harus dihadapi oleh pengajar, terutama dalam proses tahap penilaian[1]. Dalam proses ujian ada berbagai jenis soal yang digunakan, salah satunya yaitu soal Esai. Esai merupakan alat evaluasi yang memungkinkan penulisan ide dan kemampuan dalam menawarkan solusi melalui penjelasan terbuka[2]. Namun, penilaian terhadap jawaban esai sering kali dilakukan secara manual yang mengakibatkan proses memakan waktu lama dan memerlukan banyak tenaga, karena setiap esai harus diperiksa secara individual dan menimbulkan masalah lain terkait konsistensi penilai, di mana setiap penilai memiliki standar dan metode penilaian yang berbeda-beda[3]. Untuk mengatasi masalah tersebut, diperlukan sebuah sistem penilaian esai otomatis yang dapat membantu pengajar dalam memeriksa jawaban esai siswa. Hal ini memungkinkan penilaian akhir dari esai siswa dilakukan secara lebih efektif, efisien, dan objektif[4].

Automatic Essay Scoring (AES) adalah sistem yang dirancang untuk mempermudah guru dalam menilai jawaban esai peserta didik. Sistem ini beroperasi dengan cara mencocokkan kunci jawaban yang disusun oleh guru dengan jawaban yang diberikan oleh siswa[5]. Secara garis besar, cara kerja AES yaitu memberikan skor pada esai berdasarkan pada fitur yang diekstrak dari teks esai[3]. Proses penilaian melibatkan dua fase[6]. Fase pertamanya melibatkan pelatihan model untuk mengekstrak fitur-fitur yang akan digunakan sebagai kriteria penilaian esai. Sementara itu, tahap kedua melibatkan pengembangan model mesin *Automatic Essay Scoring* (AES) yang dilatih menggunakan data berlabel, berdasarkan fitur-fitur dari kumpulan data yang telah dipilih[7]. Teknologi ini memiliki peran yang sangat penting dalam mengatasi masalah yang muncul dari evaluasi manual, seperti waktu yang dibutuhkan cukup besar dan adanya subjektivitas dalam penilaian.

Penelitian ini mengimplementasikan dasar dari *Natural Language Processing* (NLP) untuk menganalisis jawaban dari siswa dan menghitung tingkat kesamaan jawaban siswa dengan kunci jawaban menggunakan algoritma Cosine Similarity. *Natural Language Processing* (NLP) adalah metode berbasis computer yang digunakan untuk memahami dan menganalisis teks atau Bahasa alami[8]. *Cosine Similarity* merupakan algoritma yang dapat digunakan untuk menilai tingkat kesamaan antara dua dokumen teks, yaitu untuk mengukur kesamaan antara kunci jawaban dan jawaban siswa[2]. Algoritma *Cosine Similarity* ini memanfaatkan kata kunci dalam sebuah dokumen sebagai ukuran untuk menghitung tingkat kesamaan antara dokumen-dokumen yang dinyatakan dalam bentuk vector[9]. Fungsi dari metode *Cosine similarity* yaitu sebagai pengukur kesamaan antara dua data berbeda dan berdasarkan pada proses perhitungan menggunakan ukuran kesamaan vektor ruang[10]. Penelitian ini menggunakan metode *N-gram* untuk memotong dan mengambil sejumlah n kata[11]. *N-gram* adalah model probabilistic yang digunakan untuk memprediksi item berikutnya dalam urutan, di mana N-gram ini diambil dari sebuah teks atau korpus dengan berbagai ukuran yaitu ukuran 1 disebut (*UNIGRAM*), ukuran 2 (*BIGRAM*), dan ukuran 3 (*TRIGRAM*)[12].

Penelitian sebelumnya yang dilakukan oleh [13] yaitu membuat suatu sistem ujian esai online dengan penilaian kemiripan jawaban menggunakan metode *Cosine Similarity* dan persamaan *Term Frequency* (TF) untuk menyamakan frekuensi setiap kata yang terdapat dalam kalimat. Untuk pengujian akurasi metode dilakukan pengujian precision, recall, dan f-measure dan berdasarkan hasil analisis dengan menggunakan metode yang telah dicoba diperoleh rata-rata 81%.

Penelitian yang pernah dilakukan oleh [14], yaitu penelitian ini mengembangkan sistem penilaian esai otomatis dengan menggunakan metode *Cosine Similarity* dan *N-gram* untuk meningkatkan efisiensi dan objektivitas dalam proses penilaian esai. *Cosine similarity* digunakan untuk mengukur kemiripan antara jawaban siswa dan kunci jawaban, sedangkan *N-gram* menganalisis pola urutan kata dalam teks. Sistem ini juga menerapkan teknik preprocessing seperti *case folding*, *tokenization*, *stopword removal*, dan *stemming* untuk mempersiapkan teks. Hasil penelitian menunjukkan bahwa sistem ini

mampu memberikan penilaian yang akurat, dengan akurasi mencapai 90%, serta meningkatkan efisiensi penilaian hingga 50% dibandingkan dengan metode manual.

Adapun penelitian terdahulu yang dilakukan[8]. Penelitian ini berhasil mengimplementasikan *Natural Language Processing* (NLP) dan algoritma Cosine Similarity dalam melakukan penilaian ujian esai secara otomatis. Hasil pengujian menunjukkan bahwa tingkat kemiripan jawaban mahasiswa terhadap kunci jawaban bervariasi berdasarkan perhitungan *Cosine Similarity*. Pada soal pertama, mahasiswa M3 memiliki tingkat kemiripan tertinggi sebesar 90,58%, sementara untuk soal kedua mahasiswa M1 mencapai tingkat kemiripan sebesar 87,71%. Adapun pada soal ketiga, mahasiswa M1 kembali memperoleh nilai tertinggi dengan tingkat kemiripan 76,70%

Penelitian lainnya yang dilakukan oleh [15], menerapkan algoritma *Levenshtein Distance*, N-Gram, dan *Synonym Replacement* dalam pengembangan sistem penilaian esai otomatis dengan harapan dapat mempermudah pengajar dalam memberikan nilai, yang sebelumnya dilakukan secara manual dan memakan waktu. Dalam implementasinya, penulis mengombinasikan algoritma *Levenshtein Distance* dengan *N-Gram* dan melakukan pengujian sebanyak empat kali, yaitu menggunakan N-Gram 1 (*Unigram*), N-Gram 2 (*Bigram*), N-Gram 3 (*Trigram*), dan N-Gram 4 (*Quadrigram*). Dataset yang digunakan dalam penelitian ini terdiri dari delapan soal dan 284 rekaman jawaban yang diambil dari hasil ujian 71 mahasiswa pada mata kuliah Bahasa Indonesia. Hasil pengujian menunjukkan bahwa sistem penilaian esai otomatis dapat mencapai akurasi tertinggi sebesar 81,7% saat menggunakan N-Gram 1.

Penelitian sebelumnya yang dilakukan oleh [16] menyoroti pentingnya penerapan sistem temu kembali informasi (*Information Retrieval System*) dalam repositori digital guna meningkatkan relevansi pencarian dokumen. Dalam penelitiannya, diterapkan metode *N-Gram* untuk mendeteksi bahasa serta *Cosine Similarity* untuk mengukur tingkat kemiripan antara query dan dokumen dalam repositori jurnal publikasi. Hasil penelitian menunjukkan bahwa sistem yang dikembangkan mampu mencapai tingkat akurasi sebesar 96,7%, presisi 85,1%, dan recall 86,9%, yang menunjukkan efektivitas pendekatan ini dalam

meningkatkan kualitas pencarian dokumen dalam repositori digital.

Pada penelitian ini, bertujuan untuk mengembangkan sistem penilaian esai otomatis berbasis *Cosine Similarity* dan *N-gram* yang efisien dan akurat. Dataset yang digunakan yaitu soal esai siswa kelas IX-A dengan jumlah siswa 19, dimana dataset ini terdiri dari 5 soal esai, jawaban siswa yang akan dianalisis untuk menentukan tingkat kesamaan dengan kunci jawaban. Selain itu, penelitian ini juga bertujuan untuk mengatasi tantangan dalam menangani variasi Bahasa dan konteks dalam jawaban esai. Penelitian ini diharapkan dapat memberikan hasil penilaian yang objektif, cepat, dan konsisten, serta mampu beradaptasi dengan berbagai pola bahasa yang digunakan oleh siswa. Dengan adanya sistem ini, diharapkan institusi pendidikan dapat menghemat waktu dalam proses evaluasi, meningkatkan kualitas umpan balik kepada siswa, serta mengurangi kemungkinan bias dalam penilaian.

1.2 Rumusan Masalah

1.2.1 Permasalahan

Berdasarkan latar belakang yang telah dijelaskan, terdapat permasalahan dalam penilaian esai seperti proses penilaian secara manual yang membutuhkan waktu lama dan menjadi beban tambahan bagi pengajar. Selain itu, subjektivitas dalam penilaian menyebabkan ketidakkonsistenan skor antara penilai, serta kesulitan dalam menangani variasi bahasa dan struktur jawaban siswa. Dalam penerapan metode *N-Gram* dan *Cosine Similarity*, terdapat beberapa tantangan, seperti pemilihan variasi *N-Gram* yang tepat untuk menangkap hubungan kata dan struktur kalimat secara optimal, kesulitan dalam menangani sinonim dan variasi struktur kalimat karena *N-Gram* hanya mempertimbangkan urutan kata tanpa memahami makna, serta sparseness data, di mana semakin tinggi nilai N, semakin jarang suatu frasa muncul sehingga dapat mempengaruhi akurasi penilaian. Selain itu, terdapat ketergantungan pada kunci jawaban, sehingga sistem perlu mampu mengenali variasi jawaban siswa agar hasil skoring lebih akurat. Oleh karena itu, diperlukan sistem Penilaian Esai Otomatis yang dapat meningkatkan objektivitas dan konsistensi dalam evaluasi jawaban esai serta mengatasi tantangan dalam penerapan metode *N-Gram* dan *Cosine Similarity*.

1.2.2 Metode Usulan

Pada penelitian ini mengusulkan sistem penilaian esai otomatis dengan metode *Cosine Similarity* untuk mengukur tingkat kesamaan antara jawaban siswa dan kunci jawaban, *N-gram* untuk menganalisis pola urutan kata dalam jawaban esai, dan *Natural Language Processing* (NLP) untuk preprocessing teks guna meningkatkan efektivitas analisis teks esai.

1.3 Pertanyaan Penelitian

Bagaimana performa metode *Cosine Similarity* dengan pendekatan *N-Gram Unigram*, *Bigram*, dan *Trigram* berdasarkan evaluasi *Mean Absolute Error* (MAE) ?

1.4 Tujuan dan Manfaat Penelitian

Tujuan dan manfaat dari penelitian yang dilakukan adalah sebagai berikut:

1.4.1 Tujuan Penelitian

1. Mengembangkan sistem penilaian esai otomatis berbasis *Cosine Similarity* dan *N-Gram* untuk meningkatkan efisiensi dan evaluasi jawaban esai.
2. Mengetahui berapa selisih nilai *Mean Absolute Error* (MAE) yang dihasilkan oleh metode *Cosine Similarity* dengan pendekatan *N-Gram Unigram*, *Bigram*, dan *Trigram*

1.4.2 Manfaat Penelitian

1. Bagi Pengajar: Memudahkan dan mempercepat proses penilaian esai, mengurangi beban kerja secara manual.
2. Bagi Institusi Pendidikan: Mendukung transformasi digital dalam sistem evaluasi pembelajaran
3. Bagi Peneliti: Memberikan kontribusi dalam pengembangan sistem *Automated Essay Scoring* (AES)

1.5 Batasan Masalah

Berdasarkan rumusan masalah yang telah diuraikan dan agar penelitian lebih terfokus, penelitian ini memiliki Batasan sebagai berikut:

1. Sistem hanya mendukung penilaian esai dalam Bahasa Indonesia.
2. Penilaian dilakukan dengan membandingkan jawaban siswa dengan kunci jawaban yang telah ditentukan oleh guru

3. Metode yang digunakan dalam sistem ini adalah *Cosine Similarity* untuk pengukuran kesamaan teks dan *N-Gram* untuk analisis pola kata.
4. N-Gram menggunakan *Unigram*, *Bigram*, dan *Trigram*

1.6 Sistematika Penulisan

Sistematika penulisan skripsi ini disusun secara sistematis dan dibagi menjadi beberapa BAB, sebagai berikut:

BAB I PENDAHULUAN

BAB ini menjelaskan latar belakang penelitian, rumusan masalah, tujuan dan manfaat penelitian, batasan masalah, serta sistematika penulisan skripsi.

BAB II TINJAUAN PUSTAKA

BAB ini membahas secara teoritis dari berbagai sumber mengenai metode *Cosine Similarity* dan *N-Gram*, dan penelitian terdahulu dari metode yang diusulkan.

BAB III METODE PENELITIAN

BAB ini menjelaskan mengenai alur penelitian yang dilakukan, arsitektur penelitian yang akan diterapkan, dataset yang digunakan, dan juga skenario uji coba.

BAB IV HASIL DAN PEMBAHASAN

BAB ini membahas hasil pengujian sistem skoring esai otomatis menggunakan metode n-gram dan *Cosine Similarity*, serta mengevaluasi performa tiap representasi n-gram melalui perhitungan Mean Absolute Error (MAE).

BAB V PENUTUP

BAB ini menyajikan kesimpulan dari seluruh penelitian yang telah dilakukan serta memberikan saran untuk pengembangan lebih lanjut berdasarkan hasil pengujian dan analisis yang telah dipaparkan pada bab sebelumnya.