

BAB I

PENDAHULUAN

1.1. Latar Belakang Penelitian

Perkembangan zaman berlangsung begitu cepat, yang menyebabkan banyaknya perubahan di berbagai aspek kehidupan. Beraneka macam inovasi diciptakan untuk memudahkan dan efisiensi pekerjaan-pekerjaan rutin manusia. Begitu pun yang seharusnya juga terjadi dalam dunia pendidikan, di mana peran teknologi dapat meringankan beban kerja guru yang cukup banyak, sehingga guru dapat fokus pada aspek-aspek interaksi edukatif pada pembelajaran di kelas. Salah satu kewajiban bagi para guru yang masih rutin dilakukan adalah proses penilaian. Sebagai contoh, penilaian hasil tugas dan jawaban ujian dalam bentuk soal uraian panjang yang dilakukan secara manual, kemudian mengunggah hasil penilaian di *e-learning* (Sanjaya, 2020). Proses penilaian yang masih dilakukan secara manual dapat menyebabkan proses penilaian yang cukup lama, pemberian nilai yang subjektivitas, akumulasi nilai yang tidak akurat, lembar penilaian yang rentan rusak dan hilang (Ningsih, 2022).

Penilaian merupakan bagian integral dari tiga pilar utama yang sangat menentukan kegiatan pembelajaran. Ketiga pilar tersebut adalah perencanaan, pelaksanaan dan penilaian. Ketika ketiga pilar ini berjalan secara terintegrasi dan saling mendukung, hal tersebut akan sangat mempengaruhi kualitas pembelajaran (Sari *et al.*, 2015). Maka dari itu penilaian harus dirancang dan dilaksanakan sesuai dengan perencanaan dan pelaksanaan pembelajaran, karena dapat menjadi tolak ukur yang akurat terhadap pencapaian kompetensi siswa berdasarkan indikator serta tujuan pembelajaran yang telah ditetapkan sebelumnya. Selain itu, penilaian juga dikaitkan dengan pengambilan keputusan atas pencapaian hasil belajar siswa setelah mengikuti proses pembelajaran (Hamir & Ekohariadi, 2022). Penilaian memiliki tujuan untuk melakukan evaluasi dari kegiatan yang sudah dilakukan agar proses pembelajaran dapat lebih baik lagi. Umumnya penilaian dibagi menjadi 3 yaitu penilaian sikap, pengetahuan, dan keterampilan. Berdasarkan kurikulum 2013, ketiga hal tersebut termasuk ke dalam penilaian autentik, yang mana bersifat satu kesatuan utuh. Adapun instrumen penilaian yang sering digunakan secara umum dibagi menjadi dua jenis yaitu tes dan non tes. Jenis tes yang biasa dilakukan

dapat berupa tes tulisan, tes lisan maupun tes tindakan. Adapun jenis tes tulisan yang sering digunakan oleh tenaga pendidik selain pilihan ganda juga dapat berupa soal uraian (esai). Ujian esai merupakan evaluasi pembelajaran yang diberikan dalam bentuk soal esai yang mempunyai jawaban lebih bervariasi dibandingkan soal pilihan ganda. Adanya variasi jawaban tersebut dapat memberikan kesulitan tersendiri bagi tenaga pendidik atau pengajar dalam menilai jawaban, baik dari segi efisiensi waktu penilaian maupun standar penilaiannya. Penilaian soal atau ujian berupa uraian yang dilakukan secara manual juga menjadi masalah bagi siswa. Salah satu masalahnya adalah siswa tidak dapat mengetahui umpan balik yang diberikan oleh guru secara detail, melainkan hanya menerima skor hasil akhir saja, tanpa diketahui alasan mengapa mereka mendapatkan nilai tersebut. Sementara itu, prinsip dasar dalam pembelajaran adalah guru memberikan umpan balik dan penguatan kepada siswa (Dimiyati, 2009).

Salah satu cara untuk menyelesaikan masalah tersebut dengan pemanfaatan teknologi *artificial intelligence* (AI), khususnya teknik *Natural Language Processing* (NLP) dan generasi bahasa alami, yang saat ini telah mengalami kemajuan yang signifikan. AI telah mengalami perkembangan baik dari segi perangkat keras maupun perangkat lunak, dan telah menghasilkan produk-produk yang digunakan dalam kehidupan sehari-hari (Putro *et al.*, 2023). Salah satu teknik NLP yang populer saat ini adalah model bahasa generatif seperti ChatGPT. Kemampuannya dapat menghasilkan teks mirip seperti manusia, berkomunikasi dengan manusia dalam berbagai konteks, mempercepat proses penulisan, menghasilkan konten yang berkualitas, dan memberikan solusi yang cepat dan akurat dari permasalahan tertentu (Misnawati, 2023). Dari kemampuan ChatGPT tersebut yang dapat memahami berbagai konteks kalimat, memungkinkan untuk membantu guru dalam melakukan penilaian terhadap variasi jawaban yang diberikan oleh siswa dalam berbagai tugas atau ujian. Lebih dari itu, ChatGPT juga memungkinkan dapat membantu guru dalam mengidentifikasi jawaban yang benar, memberikan umpan balik yang tepat, dan mengklasifikasikan tingkat pemahaman atau kemampuan dari setiap siswa. Namun ChatGPT hanya sebuah *platform* gratis dari perusahaan OpenAI untuk mendapatkan umpan balik pengguna selama masa *research preview* demi mempelajari kekuatan dan kelemahannya (OpenAI, 2022).

Untuk dapat menggunakan kemampuan model bahasa generatif di luar *platform* ChatGPT, pengguna dapat memanfaatkan *Application Programming Interface* (API) OpenAI yang disebut sebagai OpenAI GPT API. OpenAI GPT API memungkinkan pengguna mengintegrasikan kemampuan model bahasa generatif GPT secara fleksibel ke dalam aplikasi atau sistem mereka sendiri tanpa harus bergantung pada *platform* ChatGPT. Kemampuan API GPT telah teruji akurasinya dengan nilai 0.91 dalam memberikan keputusan judul dan abstrak makalah ilmiah yang termasuk atau tidak dari kriteria yang diberikan (Guo et al., 2023).

Perkembangan teknologi dapat dimanfaatkan untuk membangun sebuah sistem otomatisasi penilaian soal uraian yang efisien dan akurat. Berdasarkan tinjauan pustaka, sudah ada beberapa peneliti yang mengembangkan sistem koreksi soal esai secara otomatis. Sebagai contoh: sistem koreksi soal *essay* otomatis dengan menggunakan metode *Rabin Karp*, penilaian ini dilakukan dengan pencocokan kata atau kalimat yang sama antara jawaban siswa dengan referensi yang ada dalam sistem. Untuk mencapai hasil yang lebih baik, perlu dilakukan pembaruan dan diversifikasi data latih dalam sistem *database*. Hal ini melibatkan penambahan variasi bentuk kalimat dan pengayaan dengan sinonim kata atau makna kata yang relevan. Dengan demikian, sistem koreksi otomatis dapat memberikan hasil yang lebih akurat dan efektif (Hamza et al., 2013). Selanjutnya, sistem koreksi jawaban esai otomatis (*e-evaluation*) dengan *vector space model* pada *computer based test* (CBT), cara kerjanya mengukur nilai kemiripan dari jawaban siswa terhadap kunci jawaban (Zeniarta et al., 2020). Selain itu, sistem penilaian otomatis jawaban esai menggunakan metode GLSA, untuk penelitian lanjutan, disarankan untuk melakukan inovasi dan variasi pada beberapa teknik yang terkait dengan proses pra-pemrosesan data (Ruslan et al., 2018).

Dari ketiga penelitian terdahulu yang membahas tentang pengembangan sistem koreksi soal esai otomatis, semua cara kerja sistemnya adalah dengan melakukan komparasi atau analisis antara kunci jawaban dengan jawaban siswa. Hal tersebut menjadi faktor kekurangan dari sistem yang dikembangkan, karena cenderung membatasi kemampuan sistem untuk memahami konteks dan memberikan penilaian yang lebih holistik. Sistem yang hanya mengandalkan perbandingan antara kunci jawaban dan jawaban siswa memiliki keterbatasan

dalam mengenali nuansa, kekreatifan, dan keunikan dalam jawaban siswa yang mungkin tidak sesuai dengan kunci jawaban yang telah ditentukan. Dikarenakan dalam memeriksa soal uraian, guru menggunakan rubrik atau pedoman penskoran sebagai acuan untuk mengevaluasi pekerjaan peserta didik dan menentukan jawaban yang tepat (Musfiqon, 2016).

Oleh karena itu, peneliti bermaksud melakukan penelitian “Implementasi API OpenAI dalam Membangun Sistem Koreksi Soal Uraian Otomatis Berbasis Rubrik”. Sistem ini diharapkan menjadi solusi untuk mengatasi kekurangan yang disebutkan sebelumnya, karena menggunakan pendekatan yang lebih kompleks dengan berbasis rubrik. Berbasis rubrik yang dimaksud artinya sistem akan mengoreksi jawaban berdasarkan dengan komponen dan kriteria penilaian yang ditetapkan di setiap soal. Pada setiap komponen terdiri dari kriteria yang memiliki rentang nilai yang akan diberikan jika jawaban memenuhi kriteria dari komponen tersebut. Selain itu, pada setiap soal yang dibuat juga terdapat atribut bobot soal. Atribut ini memungkinkan guru untuk membedakan tingkat kesulitan soal. Sehingga nilai akhir yang akan diperoleh siswa dapat mencerminkan tingkat pemahaman dan penguasaan materi, mengingat bobot soal yang berbeda dapat mempengaruhi kontribusi nilai pada penilaian keseluruhan.

Dalam konteks ini, permasalahan tersebut jika dinilai dari sudut pandang *Cynefin Framework*, yang terdiri dari lima domain yaitu *Simple*, *Chaotic*, *Complicated*, *Complex*, dan *Disorder*, adalah masalah yang kompleks. Permasalahan kompleks sering kali membutuhkan pendekatan yang melibatkan interaksi dan kolaborasi antara individu yang terlibat dalam menemukan solusi. Maka dari itu, peneliti akan menggunakan metode pengembangan sistem *Prototype*. Metode *Prototype* memiliki kelebihan mendukung komunikasi yang baik antara pengembang dan pelanggan, sehingga pengembang dapat bekerja lebih baik dalam menentukan kebutuhan pelanggan (Punkastyo, 2018). Sehingga dapat menciptakan sistem penilaian yang lebih komprehensif dan objektif sesuai dengan kebutuhan guru, serta mampu mengenali dan menghargai variasi dalam jawaban siswa yang mencerminkan pemahaman yang mendalam.

1.2. Rumusan dan Batasan Masalah Penelitian

1.2.1. Rumusan Masalah

Berdasarkan pada latar belakang masalah yang telah dijelaskan, adapun rumusan masalah yang akan menjadi fokus penelitian ini adalah:

- 1) Bagaimana *prototype* sistem koreksi soal uraian otomatis berbasis rubrik?
- 2) Bagaimana implementasi API OpenAI dalam membangun sistem koreksi soal uraian otomatis berbasis rubrik?
- 3) Bagaimana nilai akurasi sistem koreksi soal uraian otomatis berbasis rubrik dalam mengoreksi jawaban yang diberikan?

1.2.2. Batasan Masalah

Mengingat keterbatasan waktu yang dimiliki dan kemampuan peneliti, maka penelitian ini perlu dibatasi untuk memudahkan dalam pemecahan masalah:

- 1) Sistem koreksi soal uraian otomatis berbasis rubrik dibangun berbasis *website* dengan bahasa PHP menggunakan *Framework Laravel*.
- 2) Implementasi API OpenAI menggunakan *runtime environment* Node.js.
- 3) Untuk menilai akurasi sistem koreksi soal uraian otomatis berbasis rubrik menggunakan metode *Mean Absolute Error (MAE)* .

1.3. Tujuan Penelitian

Berdasarkan rumusan masalah di atas, maka tujuan yang ingin dicapai dari penelitian ini antara lain:

- 1) Membangun *prototype* sistem koreksi soal uraian otomatis berbasis rubrik.
- 2) Mengimplementasi API OpenAI ke dalam sistem koreksi soal uraian otomatis berbasis rubrik.
- 3) Menilai keakuratan sistem koreksi soal uraian otomatis berbasis rubrik dalam mengoreksi jawaban yang diberikan.

1.4. Manfaat Penelitian

Adapun manfaat yang diperoleh dari penelitian implementasi API OpenAI dalam pembangunan sistem koreksi soal uraian otomatis berbasis rubrik meliputi:

1.4.1. Manfaat Teoretis

- 1) Memberikan pemahaman tentang penerapan teknologi AI, khususnya penggunaan API OpenAI, dalam penilaian otomatis soal uraian.

- 2) Memberikan wawasan tentang keterbatasan dan tantangan dalam penggunaan teknologi AI untuk penilaian otomatis, serta memberikan rekomendasi untuk pengembangan dan perbaikan metode yang lebih baik di masa depan.
- 3) Memberikan contoh nyata bagaimana teknologi AI dapat digunakan dalam proses pembelajaran dan penilaian.

1.4.2. Kebijakan

- 1) Hasil keakuratan sistem dalam memberikan penilaian, dapat digunakan sebagai dasar pengambil keputusan untuk penggunaan sistem yang lebih luas.
- 2) Memberikan acuan standar rubrik penilaian yang tepat dan dapat dipahami oleh sistem koreksi jawaban soal uraian otomatis.

1.4.3. Manfaat Praktik

- 1) Meningkatkan efisiensi dan konsistensi penilaian soal uraian.
- 2) Memberikan umpan balik yang lebih cepat kepada siswa.
- 3) Memperluas penilaian yang membutuhkan skala besar, seperti ujian sekolah.

1.4.4. Isu Sosial

- 1) Mendukung inovasi dan potensi pemanfaatan teknologi dalam pendidikan.
- 2) Memicu motivasi peneliti lain untuk melakukan pengembangan lebih lanjut.

1.5. Sistematika Penulisan

Berikut ini adalah struktur penulisan skripsi yang meliputi gambaran isi setiap bab, disusun secara terurut dari BAB I hingga BAB V, serta hubungan antara satu bab dengan bab lainnya yang membentuk kerangka skripsi.

I. PENDAHULUAN

Bab ini berisikan latar belakang penelitian yang menjelaskan konteks dan urgensi dilakukannya penelitian. Selanjutnya, terdapat rumusan dan batasan masalah penelitian yang merinci masalah-masalah yang menjadi fokus penelitian. Tujuan penelitian menjelaskan secara jelas tujuan yang ingin dicapai melalui penelitian ini. Manfaat penelitian menggambarkan nilai dari penelitian ini dalam berbagai aspek. Terakhir, sistematika penulisan skripsi menjelaskan bagaimana struktur penulisan skripsi ini disusun.

II. KAJIAN PUSTAKA

Bab ini berisikan hal-hal meliputi: Konsep-konsep, teori-teori, dalil-dalil, hukum-hukum, model-model, dan rumus-rumus utama serta turunannya dalam

bidang yang dikaji; Penelitian terdahulu yang relevan dengan bidang yang diteliti, termasuk prosedur, subjek, dan temuannya; dan Posisi teoretis peneliti yang berkenaan dengan masalah yang diteliti.

III. METODE PENELITIAN

Bab ini mencakup berbagai aspek penting dalam penelitian, seperti desain penelitian yang akan digunakan untuk mengatur kerangka penelitian ini, informasi tentang partisipan yang akan terlibat dalam penelitian, termasuk populasi dan sampel yang diambil sebagai representasi dari populasi tersebut. Selain itu, juga akan dijelaskan mengenai instrumen penelitian yang akan digunakan, prosedur penelitian yang akan dilakukan untuk mengumpulkan data, serta metode analisis data yang akan digunakan untuk menganalisis dan menginterpretasi hasil penelitian.

IV. HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil dan pembahasan dari implementasi API OpenAI dalam membangun sistem koreksi soal uraian otomatis berbasis rubrik, dengan menguraikan temuan dan analisis yang sesuai dengan tahapan metode penelitian yang digunakan.

V. SIMPULAN, IMPLIKASI, DAN REKOMENDASI

Bab ini berisikan kesimpulan, implikasi, dan rekomendasi berdasarkan hasil analisis temuan penelitian, serta memberikan penafsiran dan makna yang diberikan oleh peneliti terhadap hasil tersebut. Bab ini juga mengungkapkan hal-hal penting yang dapat diambil manfaatnya dari hasil penelitian.