

BAB I

PENDAHULUAN

Bab pendahuluan ini membahas mengenai latar belakang, rumusan masalah, tujuan dan manfaat penelitian, ruang lingkup penelitian, serta sistematika penulisan yang digunakan dalam skripsi yang berjudul “*Automatic Feedback Untuk Jawaban Soal Uraian Singkat Berbahasa Indonesia Menggunakan Arsitektur DistilBART*”.

1.1 Latar Belakang

Evaluasi merupakan bagian yang sangat penting dalam Pendidikan karena berfungsi untuk mengukur efektivitas pembelajaran serta tingkat pemahaman mahasiswa terhadap materi yang telah diajarkan. Evaluasi bertujuan untuk mengetahui tingkat keberhasilan peserta didik dalam memahami pengetahuan yang diberikan oleh pendidik serta menjadi alat ukur untuk meningkatkan metode pengajaran yang lebih efektif (Musarwan & Warsah, 2022). Menurut Idrus, (2019), evaluasi pembelajaran memiliki beberapa fungsi utama, seperti alat ukur keberhasilan pembelajaran, alat pengembangan pembelajaran, dan alat peningkatan motivasi mahasiswa. Selain itu, evaluasi bukan hanya menilai hasil belajar, tetapi juga mencakup proses pembelajaran itu sendiri, termasuk strategi pengajaran, pemahaman mahasiswa, dan kualitas *feedback* yang diberikan oleh pendidik (Musarwan & Warsah, 2022). Dalam praktiknya, terdapat berbagai metode evaluasi yang digunakan, salah satunya melalui soal uraian (esai) dan soal jawaban singkat. Soal jawaban singkat (*short answer questions*) banyak digunakan dalam evaluasi pembelajaran karena kemampuannya dalam menilai pemahaman konseptual mahasiswa secara lebih mendalam dibandingkan soal pilihan ganda (Condor & Pardos, 2022). Selain itu, penggunaan soal jawaban singkat dapat meningkatkan keterampilan berpikir kritis karena mahasiswa harus mengorganisir jawabannya secara ringkas namun tetap mencakup poin-poin penting dari materi yang dipelajari.

Penggunaan soal uraian dalam evaluasi pembelajaran menghadirkan tantangan tersendiri karena proses penilaiannya memerlukan keterlibatan mendalam dari pendidik untuk membaca, memahami, dan memberikan *feedback* yang tepat terhadap setiap jawaban mahasiswa, seperti yang diungkapkan oleh Henderson dkk., (2019). *Feedback* yang efektif

tidak hanya membantu mahasiswa dalam mengidentifikasi kesalahan dan meningkatkan motivasi belajar, tetapi juga mendukung pengembangan keterampilan berpikir kritis melalui proses refleksi dan revisi (Wisniewski dkk., 2020). Selain itu, Panadero dkk., (2018) menekankan bahwa integrasi antara penilaian formatif dan pembelajaran mandiri dapat memperkuat regulasi diri mahasiswa, sementara Imran dkk., (2024) menunjukkan bahwa penerapan teknologi kecerdasan buatan dalam penilaian menawarkan solusi inovatif untuk pemberian *feedback* yang efisien, sehingga secara sinergis dapat meningkatkan kualitas pembelajaran secara keseluruhan.

Dengan berkembangnya teknologi, kecerdasan buatan (*Artificial Intelligence / AI*) dan *Natural Language Processing (NLP)* mulai diterapkan dalam berbagai bidang, termasuk dalam evaluasi pembelajaran. *NLP*, yang merupakan cabang dari *AI*, memungkinkan mesin dalam memahami, mengolah, dan menganalisis bahasa alami dalam interaksi manusia. Dalam konteks evaluasi pendidikan, *NLP* berperan dalam mengotomatisasi proses penilaian soal uraian dan jawaban singkat, sehingga dapat mengurangi beban kerja pendidik serta meningkatkan efisiensi dan objektivitas evaluasi (Sihombing, 2022). Untuk itu diperlukan model yang dapat menghasilkan *feedback* yang memberikan informasi terhadap kesalahan pada jawaban mahasiswa serta dapat meningkatkan efisiensi dan objektivitas pada evaluasi secara otomatis. Pada penelitian terdahulu yang dilakukan oleh Koe dkk., (2024), untuk mengembangkan sistem *automatic feedback* berbasis *Artificial Intelligence (AI)* dengan teknologi *Natural Language Processing (NLP)* dan model *OpenAI GPT* dalam pembelajaran daring. Sistem ini dirancang untuk menganalisis respons mahasiswa secara otomatis, memberikan *feedback* yang cepat dan relevan. Selanjutnya, penelitian oleh Guo dkk., (2024), melakukan eksplorasi penggunaan *multi-agent system* dalam meningkatkan kualitas *feedback* yang dihasilkan oleh *Generative AI*. Sistem yang dikembangkan ini terdiri dari dua agen *AI*, yaitu agen pertama yang bertugas menghasilkan *feedback* dan agen kedua yang berfungsi sebagai validator yang memperbaiki *feedback*. Oleh karena itu, dalam penelitian ini menggunakan *text generation* sebagai pendekatan untuk mengembangkan model *automatic feedback*. Pada penelitian terdahulu yang dilakukan oleh (Tangsali dkk., 2022), penelitian ini melakukan perbandingan terhadap beberapa arsitektur *Transformer* dalam tugas *abstractive summarization*, salah satunya *BART-Large*. *BART-Large* mengadopsi *transformer encoder-decoder (seq2seq)*, yang telah terbukti efektif dalam berbagai tugas *text generation*.

Akan tetapi, *BART-Large* merupakan model besar yang kompleks dengan arsitektur *encoder-decoder* yang terdiri dari 12 *layer* (Lewis dkk., 2019). *BART-Large* secara khusus dilatih pada *dataset* berskala besar dan memiliki ukuran model yang lebih besar dibandingkan dengan model *BART base* (Mercan dkk., 2023). Oleh karena itu, arsitektur *DistilBART* akan digunakan sebagai pendekatan utama dalam mengembangkan model untuk menghasilkan *feedback* secara otomatis pada soal uraian berbahasa Indonesia. *DistilBART* merupakan versi lebih ringan dari *BART* yang telah melalui proses distilasi untuk mengurangi ukuran model tanpa mengorbankan banyak kinerja (Shleifer & Rush, 2020). Arsitektur *DistilBART* memiliki keunggulan dalam efisiensi pemrosesan teks, yang sangat penting dalam tugas-tugas yang membutuhkan analisis bahasa alami, seperti pembuatan rangkuman dan analisis teks (Lewis dkk., 2019), *DistilBART* dipilih juga karena jumlah *dataset* yang digunakan terbatas, sehingga penggunaan arsitektur yang lebih kecil dan efisien menjadi solusi yang tepat untuk menghindari *overfitting* dan memastikan model tetap dapat melakukan generalisasi dengan baik pada data baru. Selain itu, penelitian oleh Sanh dkk., (2019), menunjukkan bahwa model hasil distilasi dapat mempertahankan kinerja yang mendekati model aslinya dengan kebutuhan komputasi yang lebih rendah, sehingga lebih efisien untuk diterapkan dalam skenario dengan sumber daya yang terbatas.

Selain menggunakan *DistilBART*, dalam pengembangan model *automatic feedback* akan memanfaatkan metode *fuzzy matching* berbasis *Partial Ratio* untuk menghitung tingkat kesamaan antara jawaban mahasiswa dan jawaban referensi. Studi yang dilakukan oleh (Rao dkk., 2022), menunjukkan bahwa *Partial Ratio* dapat digunakan secara efektif dalam mendeteksi kemiripan teks dalam berbagai bahasa, termasuk dalam sistem deteksi plagiarisme. Oleh karena itu, diusulkan penerapan *DistilBART* untuk pengembangan model *Automatic Feedback* pada *dataset* uraian singkat berbahasa Indonesia.

Dengan dikembangkannya model *feedback* otomatis berbasis *NLP*, diharapkan dapat membantu meningkatkan kualitas evaluasi pembelajaran dengan memberikan *feedback* yang lebih mendalam, relevan, dan adaptif terhadap jawaban mahasiswa. Model juga diharapkan dapat mengurangi beban kerja pendidik serta meningkatkan pengalaman belajar mahasiswa dengan menyediakan *feedback* yang informatif dan konstruktif dalam pembelajaran berbasis soal uraian.

1.2 Rumusan Masalah

Berdasarkan latar belakang sebelumnya, rumusan masalah yang dibahas pada penelitian ini adalah bagaimana mengembangkan model *Automatic Feedback* menggunakan pada jawaban soal uraian singkat berbahasa Indonesia menggunakan *DistilBART*. Dari permasalahan tersebut diturunkan ke dalam permasalahan khusus yaitu:

1. Bagaimana menentukan konfigurasi *hyperparameter* terbaik terkhususnya *batch size* dan *learning rate* untuk *fine-tuning* model *DistilBART* dan *BART-Large* agar mendapatkan hasil kinerja yang optimal untuk *automatic feedback* pada soal uraian singkat?
2. Bagaimana perbandingan hasil kinerja dan evaluasi dari *DistilBART* dengan *BART-Large* dalam menghasilkan *feedback* secara otomatis pada soal uraian singkat?

1.3 Tujuan dan Manfaat

Dalam mengembangkan model terdapat beberapa tujuan yang hendak dicapai. Adapun tujuan yang dimaksud antara lain:

1. Menghasilkan konfigurasi konfigurasi *hyperparameter* terbaik terkhususnya *batch size* dan *learning rate* untuk *fine-tuning* model agar mendapatkan hasil kinerja yang optimal untuk *automatic feedback* pada soal uraian singkat.
2. Mendapatkan selisih perbandingan hasil kinerja dan evaluasi dari model *DistilBART* dan *BART-Large* dalam menghasilkan *feedback* secara otomatis pada soal uraian singkat.

Adapun manfaat yang diharapkan dapat diperoleh dari penelitian ini adalah mendapatkan konfigurasi konfigurasi *hyperparameter* terbaik terkhususnya *batch size* dan *learning rate* untuk melakukan *fine-tuning* pada *DistilBART* dan *BART-Large*, serta mengetahui kinerja terbaik antara *DistilBART* dan *BART-Large* dalam menghasilkan *feedback* secara otomatis pada soal uraian singkat. Sehingga, model dapat digunakan untuk menghasilkan *feedback* yang informatif yang bermanfaat untuk memahami kesalahan pada jawaban mahasiswa secara otomatis dan menjadi referensi bagi penelitian selanjutnya dalam mengembangkan sistem *automatic feedback* yang lebih akurat dan efisien.

1.4 Ruang Lingkup

Ruang lingkup penelitian ini mencakup beberapa aspek utama yang menjadi batasan dalam pelaksanaan penelitian. Hal ini untuk memfokuskan penelitian pada aspek-aspek yang relevan dengan tujuan yang ingin dicapai. Adapun ruang lingkup penelitian ini yaitu optimasi meliputi pengujian berbagai nilai untuk *threshold similarity*, *learning rate*, *batch size* yang digunakan dalam model untuk setiap *dataset*.

1.5 Sistematika Penulisan

Sistematika penulisan digunakan untuk memberikan Gambaran dan alur yang runtut terkait penyusunan laporan skripsi yang berjudul *Automation Feedback* pada Soal Uraian Berbahasa Indonesia Menggunakan Arsitektur *DistilBART*. Sistematika penulisan laporan dijelaskan sebagai berikut:

BAB I PENDAHULUAN

Bab pendahuluan menyajikan latar belakang masalah, rumusan masalah, tujuan dan manfaat penelitian, ruang lingkup penelitian, serta sistematika penulisan laporan yang dibuat.

BAB II TINJAUAN PUSTAKA

Bab tinjauan pustaka memuat teori-teori yang digunakan pada penelitian ini. Bab ini mengulas beberapa terminologi antara lain *state of the art*, pra-pemrosesan data, arsitektur *transformer*, arsitektur *bart*, *bart-large*, *distilbart*, dan evaluasi model *automatic feedback*.

BAB III METODOLOGI PENELITIAN

Bab ini menyajikan alur dan langkah-langkah yang digunakan sebagai acuan dalam pelaksanaan penelitian. Gambaran umum dari tahap penelitian ini adalah pengumpulan data, pra-pemrosesan data, tokenisasi data, pelatihan model *automatic feedback*, evaluasi model *automatic feedback*, dan eksperimental setup.

BAB IV HASIL DAN PEMBAHASAN

Bab ini membahas mengenai pembahasan dari seluruh langkah penelitian dan penjabaran hasil dan analisis eksperimen yang didapatkan dari pelatihan model *automatic feedback* menggunakan arsitektur *DistilBART* dan *BART-Large* sesuai dengan metodologi penelitian yang dijadikan acuan.

BAB V PENUTUP

Bab penutup berisi Kesimpulan terhadap hasil dan analisis dari eksperimen yang telah dilakukan pada penelitian ini terkait pengembangan model *Automatic Feedback* dan saran yang dibutuhkan untuk penelitian selanjutnya.