BAB I
PENDAHULUAN

Bab pendahuluan ini membahas mengenai latar belakang masalah, rumusan masalah, tujuan dan manfaat, ruang lingkup, serta sistematika penulisan yang digunakan dalam dokumen skripsi ini yang berjudul Klasifikasi Jawaban *Generative AI* Versus Jawaban Manusia Pada Teks Bahasa Indonesia Menggunakan Metode TF-IDF dan *Support Vector Machine.*

1.1 Latar Belakang

Artificial Intelligence (AI) telah mengalami perkembangan pesat dan memiliki dampak signifikan di berbagai sektor dalam dekade terakhir. Penggunaan AI tidak hanya terbatas pada industri telekomunikasi, tetapi juga telah merambah ke sektor-sektor lain seperti perbankan, manufaktur, jasa, pemerintahan, dan bahkan pendidikan. AI telah menjadi bagian integral dari perkembangan teknologi informasi dan komunikasi, membawa perubahan besar dalam berbagai aspek kehidupan manusia. Penerapan AI dalam berbagai bidang menunjukkan betapa pentingnya peran teknologi AI dalam mendorong inovasi dan efisiensi di era digital saat ini (Ririh dkk., 2020; Rifky, 2024).

Teknologi AI terus mengalami perkembangan yang luar biasa (Wamba dkk., 2021). Beberapa teknologi yang termasuk di dalamnya adalah *machine learning*, *deep learning*, *voice recognition*, dan *natural language processing* (NLP). Salah satu inovasi AI yang mendapat banyak perhatian adalah *Large Language Model* (LLM), yaitu model bahasa besar hasil dari kemajuan *neural network* (NN) dan NLP, yang dikenal dengan *Language Model* (LM). LLM merupakan model *Generative AI* yang mampu memproses dan menghasilkan teks dalam bahasa manusia. Model LLM dilatih menggunakan data teks yang sangat besar, dengan total ukuran mencapai lebih dari 774,5 *terabyte* untuk *pre-training corpora* dan lebih dari 700 juta *instance* untuk jenis *dataset* lainnya. Data ini mencakup 444 dataset dari 8 kategori bahasa dan 32 domain, yang masing-masing digolongkan ke dalam 20 dimensi berbeda (Liu dkk., 2024). Selain itu, model tersebut menggunakan teknik pembelajaran mendalam untuk memahami pola dan struktur bahasa (Hadi dkk., 2023). LLM memungkinkan berbagai

aplikasi bahasa, seperti kemampuan dalam pembuatan teks, penerjemahan, penyusunan ringkasan, menjawab pertanyaan, menulis kode, dan analisis sentimen (Hadi dkk., 2023). Namun, kemampuan tersebut menimbulkan tantangan baru, terutama dalam hal membedakan teks yang dihasilkan oleh AI dengan teks yang ditulis oleh manusia.

Tantangan penggunaan AI semakin nyata, terutama dalam konteks pendidikan. Banyak peserta didik memanfaatkan alat berbasis *Generative AI*, seperti ChatGPT (*Generative Pre-training Transformer*), sebagai alternatif untuk menyelesaikan tugas yang diberikan oleh guru atau dosen. Bahkan, tidak sedikit yang menggunakan ChatGPT dalam penyusunan skripsi guna memenuhi persyaratan tugas akhir mereka (Hidayanti & Azmiyanti, 2023). Popularitas ChatGPT yang terus meningkat turut memperkuat alasan penggunaannya, di mana *chatbot* AI, termasuk ChatGPT, mencatatkan 14,6 miliar kunjungan antara September 2022 hingga Agustus 2024 (Kedaton, 2024). Kondisi tersebut menimbulkan kekhawatiran di kalangan pendidik, yang merasa bahwa ChatGPT dapat berdampak negatif, seperti mengancam dan merusak kompetensi akademik peserta didik, baik di sekolah maupun perguruan tinggi. Selain itu, dampak negatif lainnya yang ditimbulkan adalah ketergantungan pada teknologi, bias dan ketidakakuratan data, serta masalah privasi dan keamanan data (Bukhori dkk., 2024). Jawaban yang dihasilkan oleh ChatGPT sering kali menyerupai respons manusia (Wahid dkk., 2023).

Kemampuan membedakan jawaban yang dihasilkan oleh *Generative AI* dan jawaban manusia semakin penting, terutama dalam konteks pendidikan. Menurut Campino (2024), kemampuan tersebut sangat diperlukan untuk memastikan bahwa AI digunakan secara etis dalam pendidikan, meningkatkan kesadaran untuk mencegah penyalahgunaan, serta mendorong peserta didik mengembangkan keterampilan penalaran mereka sendiri. Penilaian akademik, seperti tugas, kuis, atau ujian, harus memastikan bahwa jawaban mencerminkan pemahaman dan kemampuan analitis pelajar, bukan sekadar hasil dari AI. Hal tersebut penting untuk mengidentifikasi kekuatan dan kelemahan peserta didik serta membantu pendidik menilai efektivitas strategi pembelajaran yang digunakan (Ermawati & Hidayat, 2017). Jawaban AI adalah respons yang dihasilkan oleh model *Generative AI* seperti ChatGPT, yang merespons pertanyaan berdasarkan pola data besar yang telah dilatih untuk

menghasilkan teks yang koheren dan relevan. Sebaliknya, jawaban manusia mencerminkan pemahaman kontekstual, pengalaman, dan perspektif pribadi yang diperoleh melalui proses pembelajaran. Oleh karena itu, diperlukan metode yang akurat untuk mengklasifikasikan jawaban *Generative AI* dan jawaban manusia dalam teks berbahasa indonesia.

Pada era perkembangan teknologi yang semakin pesat, teknik klasifikasi, sebagaimana dikemukakan oleh Priyambodo dan Prihati (2020), telah menjadi salah satu metode paling esensial dalam berbagai aplikasi *machine learning*. Klasifikasi tidak hanya berfungsi sebagai alat bantu dalam pengelompokan data, tetapi juga memegang peran penting dalam proses analisis data yang lebih kompleks, seperti penambangan data dan teks. Teknik klasifikasi dalam klasifikasi teks memungkinkan pengelompokan informasi secara efektif ke dalam kategori yang telah ditentukan. Hal ini berguna dalam berbagai bidang, mulai dari analisis sentimen hingga penyortiran dokumen otomatis. Proses klasifikasi yang memanfaatkan data pelatihan berlabel memberikan kerangka kerja yang jelas dan sistematis untuk membangun model yang mampu mengkategorikan data baru dengan akurasi yang tinggi. Proses klasifikasi menurut Nicolosi (2008) meliputi tahap pembelajaran yang menganalisis data pelatihan dan membentuk aturan klasifikasi serta tahap klasifikasi yang menggunakan aturan tersebut untuk mengelompokkan data uji ke dalam kategori yang sesuai. Dengan demikian, klasifikasi berperan tidak hanya sebagai alat teknis, tetapi juga sebagai bagian penting dari strategi analisis data yang lebih besar, memungkinkan pengambilan keputusan berbasis data yang lebih cerdas dan efisien.

Penelitian terdahulu menunjukkan bahwa berbagai metode klasifikasi teks telah diterapkan dalam berbagai domain teks. Hayawi dkk. (2023) menganalisis teks berbahasa Inggris seperti esai dan puisi menggunakan algoritma *machine learning* klasik seperti *Random Forest*, *Support Vector Machine* (SVM), *Logistic Regression*, serta algoritma *deep learning* LSTM. Dari hasil penelitian tersebut, SVM berhasil mencapai akurasi tertinggi sebesar 99,85% dalam membedakan teks manusia dari GPT. Di sisi lain, Hermanto dkk. (2020) melakukan klasifikasi komplain mahasiswa dengan algoritma SVM dan *Naïve Bayes*, di mana SVM mencapai akurasi 84,45%, lebih tinggi dibandingkan *Naïve Bayes* yang hanya 69,75%. Serupa dengan hal tersebut, Sandag dkk. (2018) mengklasifikasikan SMS spam dengan algoritma yang

sama, menggunakan data dari *database* Kaggle. Setelah melalui proses *tokenizing*, *normalize*, *filtering*, dan *stemming*, hasilnya menunjukkan bahwa SVM memiliki performa lebih baik dengan akurasi 96,72%, dibandingkan *Naïve Bayes* yang hanya 81,77%.

Penelitian lain juga menegaskan keunggulan SVM dalam berbagai kasus klasifikasi. Muis dan Affandes (2015), misalnya, metode SVM berhasil mengklasifikasikan tweet menggunakan kernel *Radial Basis Function* (RBF), mencapai akurasi tertinggi 99,12% pada data yang telah melalui proses pemilihan fitur. Ropikoh dkk. (2021) dalam penelitiannya tentang klasifikasi berita *hoax* COVID-19 membandingkan dua jenis kernel, yaitu kernel *linear* dan RBF, dan menemukan bahwa SVM dengan kernel *linear* pada skenario pembagian data 80:20 mencapai akurasi 97,06%, sedangkan kernel RBF pada skenario 90:10 hanya mencapai 90,46%. Selain itu, Maulina dan Sagara (2018) menggunakan SVM dengan pembobotan *Term Frequency-Inverse Document Frequency* (TF-IDF) untuk mengklasifikasikan artikel *hoax* dan berhasil mencapai akurasi 95,83%. Penelitian Rahman dkk. (2021) juga mendukung temuan tersebut, di mana SVM dengan kernel RBF berhasil mengklasifikasikan ujaran kebencian di Twitter dengan akurasi 93% pada 700 data latih dan 300 data uji.

Beberapa penelitian menunjukkan bahwa metode TF-IDF terbukti efektif dalam representasi fitur dasar yang cukup kuat untuk membedakan kata-kata dengan bobot penting tanpa memerlukan sumber daya komputasi yang besar. Sementara itu, SVM tetap menjadi salah satu metode utama dalam klasifikasi teks, terutama terbukti *roboust* serta fleksibilitasnya dalam menangani data *linear* dan *non-linear* melalui variasi kernel. Penerapan metode TF-IDF dan SVM telah banyak dilakukan di berbagai bidang, namun penelitian ini secara khusus berfokus pada klasifikasi teks antara jawaban *Generative AI* versus jawaban manusia dalam bahasa Indonesia, suatu area yang masih jarang dieksplorasi. Oleh karena itu, penelitian ini bertujuan mengisi kekosongan tersebut sekaligus berkontribusi pada pengembangan teknologi pemrosesan bahasa alami di Indonesia. Untuk mencapai tujuan tersebut, penulis mengusulkan penggunaan metode TF-IDF dan SVM dalam Klasifikasi Jawaban *Generative AI* Versus Jawaban Manusia.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah disusun, rumusan masalah dalam penelitian ini adalah bagaimana performa kombinasi metode TF-IDF sebagai teknik ekstraksi fitur dan SVM dengan berbagai pilihan kernel dalam mengklasifikasikan jawaban *Generative AI* Versus Jawaban Manusia untuk memperoleh akurasi yang optimal.

1.3 Tujuan dan Manfaat

Tujuan dari penelitian ini adalah mendapatkan performa model klasifikasi terbaik dengan mengombinasikan TF-IDF sebagai teknik ekstraksi fitur dan metode SVM, serta memilih parameter optimal untuk mendapatkan model terbaik pada setiap kernel, sehingga dapat diidentifikasi kernel SVM yang paling efektif dalam mengklasifikasikan jawaban *Generative AI* Versus Jawaban Manusia.

Manfaat dari penelitian ini adalah tersedianya model klasifikasi yang mampu membedakan jawaban *Generative AI* Versus Jawaban Manusia menggunakan representasi fitur TF-IDF dan metode SVM dengan pemilihan kernel terbaik. Model tersebut diharapkan dapat digunakan untuk pengembangan sistem klasifikasi otomatis yang lebih akurat di masa mendatang, terutama dalam bidang seperti klasifikasi teks, pendidikan, dan pemrosesan bahasa alami.

1.4 Ruang Lingkup

Penelitian skripsi ini diperlukan adanya ruang lingkup agar skripsi yang dilakukan lebih terarah dan mencapai sasaran yang diharapkan. Ruang lingkup dalam skripsi ini adalah sebagai berikut:

1. Pengumpulan data dilakukan dengan pemberian kuis mata kuliah *Internet of Things* (IoT) kepada mahasiswa Informatika Semester II Universitas Diponegoro melalui platform *e-learning* Universitas Diponegoro (Kulon2 Undip) dan mengunduh jawaban dari *platfrom* tersebut dalam format *comma-separated values* (CSV). Sementara itu, jawaban dari AI diperoleh dengan mengajukan soal kuis menggunakan beberapa platform *Generative AI* (ChatGPT-4o mini, Gemini 1.0, Bing AI, dan Perplexity) secara manual. Jawaban *Generative AI* dikumpulkan dan diformat menjadi CSV.

2. Kuis ini terdiri dari tiga soal, namun soal kedua dan ketiga yang akan digunakan untuk analisis. Soal kedua meminta menyebutkan nama komponen utama dalam arsitektur komputer beserta fungsinya, sedangkan soal ketiga meminta menjelaskan tentang perbedaan antara mikrokontroler dan komputer personal.
3. Rentang waktu pengumpulan data dari kuis tersebut adalah dari tanggal 29 April 2024 hingga tanggal 20 Juni 2024.
4. Rentang waktu pengumpulan data dari berbagai platfrom *Generative AI* adalah dari tanggal 8 Oktober 2024 hingga tanggal 10 Oktober 2024.
5. Data yang terkumpul terdiri dari 411 baris jawaban *generative AI* dan 408 baris jawaban mahasiswa.
6. Data yang terkumpul pada penelitian ini belum melalui proses validasi penuh untuk memastikan bahwa jawaban manusia benar-benar bebas dari bantuan *generative AI*.

1.5 Sistematika Penulisan

Struktur penulisan dalam laporan ini dibagi menjadi lima bab utama yang membahas penelitian "Klasifikasi Jawaban *Generative AI* Versus Jawaban Manusia Pada Teks Bahasa Indonesia Menggunakan Metode TF-IDF dan *Support Vector Machine*". Pembagian ini disusun dengan tujuan agar penulisan menjadi lebih sistematis dan mudah dipahami. Berikut ini adalah ringkasan singkat dari setiap bab yang dibahas:

BAB I PENDAHULUAN

Bab ini menyajikan latar belakang masalah, rumusan masalah, tujuan dan manfaat, ruang lingkup, serta sistematika penulisan skripsi.

BAB II TINJAUAN PUSTAKA

Bab ini menyajikan studi pustaka mengenai dasar teori yang digunakan dalam penelitian skripsi ini: *state-of-the-art*, *Artificial Intelligence* (AI), *Natural Language Processing* (NLP), *Large Language Models*, *Machine Learning* (ML), klasifikasi teks, *text preprocessing*, *Bag-of-Words* (BoW), *Term Weighting*, *K-Fold Cross Validation, Support Vector Machine* (SVM), evaluasi model, serta *tools* dan *library*.

BAB III METODE PENELITIAN

Bab ini menjelaskan metodologi yang digunakan dalam penelitian ini, meliputi alur penelitian, proses pengumpulan dan pelabelan data, pemrosesan data, pembagian *dataset*, serta pembobotan data menggunakan metode TF-IDF. *K-Fold Cross Validation* digunakan pada data latih untuk memastikan pemilihan parameter optimal pada model SVM. Selain itu, dijelaskan juga pembangunan model klasifikasi menggunakan algoritma SVM dengan berbagai jenis kernel, yaitu *linear*, RBF, dan *polynomial*. Pada bagian akhir, dilakukan pengujian dan evaluasi model klasifikasi untuk membedakan antara jawaban yang diberikan oleh *Generative AI* versus jawaban manusia menggunakan SVM.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menjelaskan analisis dan hasil skripsi, serta membahas implementasi pelatihan dan skenario pengujian menggunakan tiga jenis kernel: *linear*, RBF, dan *polynomial*. Model dilatih menggunakan *K-Fold Cross Validation* untuk mencari parameter optimal melalui g*rid search*. Parameter optimal yang diperoleh dari proses ini kemudian digunakan untuk pengujian akhir pada data uji.

BAB V PENUTUP

Bab ini berisi kesimpulan dari bab-bab yang telah dijelaskan sebelumnya dan saran untuk pengembangan penelitian lebih lanjut.