

BAB 1 PENDAHULUAN

1.1 Latar Belakang Masalah

Kecerdasan buatan atau *Artificial Intelligence (AI)* merupakan sistem *Machine-Based* yang mampu membuat sebuah rekomendasi, prediksi, dan diagnosa yang sama seperti manusia[4]. Dalam beberapa tahun belakang, dampak dari kemajuan yang terjadi pada *AI* telah memasuki banyak bidang seperti bisnis, agrikultur, kesehatan, manufaktur, pendidikan, seni, dan bidang lainnya[5]. Meskipun begitu, *AI* dianggap dapat mengancam penyaluran tenaga kerja dikarenakan kompleksitas pekerjaan yang mulai bisa diambil alih oleh kecerdasan buatan demi efisiensi pekerjaan[6]. Walaupun begitu, masyarakat Indonesia memiliki pandangan yang berbeda terhadap kecepatan inovasi *AI* seringkali ditemukan di berbagai kolom komentar platform media sosial seperti *Twitter*.

AI-generated artwork merupakan setiap bentuk karya seni digital yang dihasilkan melalui bantuan kecerdasan buatan, di mana proses penciptaannya melibatkan algoritma atau model komputasional tertentu yang dirancang untuk meniru, mereplikasi, atau bahkan menciptakan gaya artistik layaknya manusia. Teknologi ini memungkinkan sistem *AI* untuk memproses masukan berupa teks deskriptif, referensi visual, atau parameter teknis lainnya, kemudian menghasilkan gambar, ilustrasi, atau karya seni secara otomatis tanpa memerlukan keterlibatan langsung dari seniman dalam proses menggambar atau melukis. Penggunaan *AI* dalam pembuatan karya seni membawa banyak dampak positif dan negatif terhadap perkembangan teknologi, dengan plagiarisme dan pemanfaatan *AI* yang berkesan tidak bermoral atau beretika, serta melanggar hak cipta dari seniman yang karyanya digunakan tanpa izin dalam proses pembuatan karya digital[7].

Salah satu contoh dari program aplikasi yang digunakan untuk membuat karya digital dengan bantuan proses dari kecerdasan buatan adalah *Midjourney*. *Midjourney* merupakan program kecerdasan buatan generatif yang digunakan melalui ketikan berdeskriptif[8]. Berdasarkan data statistik yang diberikan oleh *Midjourney*, mereka telah mendapatkan 14 juta pengguna yang telah meregistrasikan akun untuk menggunakan *Midjourney*[9]. Setiap pengguna memiliki tujuan masing-masing dalam menggunakan kecerdasan buatan untuk menghasilkan karya digital secara efisien tanpa memerlukan pengalaman atau

keahlian artistik tertentu, dan semua ini didorong oleh keinginan untuk mengekspresikan diri mereka secara instan[10]. Namun, dalam lima tahun terakhir, muncul berbagai kontroversi yang signifikan terkait penyalahgunaan AI-generated artwork, seperti kasus di mana karya seni yang dihasilkan oleh AI, yang diduga dilatih menggunakan dataset tanpa izin dari seniman manusia, berhasil memenangkan kompetisi seni konvensional dan memicu protes luas dari komunitas kreatif[10][11]. Salah satu insiden yang mencuat terjadi pada tahun 2022 ketika sebuah karya AI memenangkan penghargaan seni di Colorado State Fair dan memicu perdebatan etis terkait plagiarisme dan validitas artistik karya tersebut[11]. Hal ini menyoroti bagaimana penggunaan AI dalam seni dapat melanggar hak cipta, menghapus nilai orisinalitas, dan merugikan seniman yang karyanya dijadikan data latih tanpa persetujuan eksplisit.

Oleh karena itu, untuk memahami lebih dalam bagaimana masyarakat merespons tren ini, diperlukan suatu pendekatan sistematis untuk mengukur dan menganalisis opini publik. Salah satu metode yang dapat digunakan adalah analisis sentimen, yaitu proses untuk mengungkap dan mengkategorikan pendapat yang diungkapkan dalam bentuk teks guna menentukan apakah suatu tanggapan bersifat positif, negatif, atau netral[12]. Teknik yang digunakan dalam melakukan analisis sentimen mengandalkan pembelajaran mesin dalam mengidentifikasi suatu teks untuk mengambil kata-kata kunci atau ciri-ciri terkait dalam mengelompokkan teks ke dalam kategori sentimen yang sesuai[13]. Analisis Sentimen dapat diterapkan di berbagai macam media sosial, situs web, atau data lain berisi teks[13] yang dapat dibaca oleh komputer.

Dalam konteks ini, Twitter dipilih sebagai sumber data karena tingginya intensitas diskusi publik mengenai AI-generated artwork di platform tersebut. Twitter dipilih sebagai sumber data karena merupakan platform mikroblog dengan API streaming yang mendukung pengambilan data real-time dan merupakan gudang opini warga dunia[14]. Berdasarkan pengamatan awal, topik ini sering menimbulkan respon emosional dari pengguna, mulai dari apresiasi terhadap teknologi hingga kekhawatiran terhadap dampak sosial dan etisnya. Hal ini menjadikan Twitter sebagai medium yang relevan untuk dianalisis menggunakan pendekatan NLP, khususnya dalam mengklasifikasikan opini berdasarkan sentimen yang terkandung di dalamnya[14].

Suatu bahasa dapat didefinisikan sebagai suatu rangkaian peraturan dan pasangan simbol. Simbol akan digabungkan dan digunakan untuk menyampaikan atau menyebarkan informasi tersebut [15]. *Natural Language Processing (NLP)*

merupakan cabang ilmu *Artificial Intelligence* yang digunakan untuk memahami dan mereplikasikan bahasa manusia melalui teks dan/atau kata yang terucapkan layaknya seorang manusia [15] [16]. *NLP* biasa dikaitkan dengan berbagai teori dan teknik dalam menghadapi masalah dari berkomunikasi dengan komputer menggunakan bahasa manusia, dan terdiri dari dua komponen, yaitu *Natural Language Understanding (Linguistic) (NLU)*, dan *Natural Language Generator (NLG)* [15].

IndoNLP (Indonesia *Natural Language Processing*) merupakan *benchmark* Indonesia pertama untuk *Natural Language Understanding (NLU)* yang terdiri atas kumpulan dua belas tugas yang beragam. Tugas utama dikategorikan berdasarkan input, seperti kalimat tunggal dan pasangan kalimat, dan tujuan, seperti tugas klasifikasi kalimat dan pelabelan urutan tugas, *Benchmark* ini dirancang untuk memenuhi berbagai gaya dalam bahasa Indonesia formal dan sehari-hari yang sangat beragam[17]. *IndoNLP* menyediakan berbagai macam model *NLU* yang menggunakan pretrained BERT dengan Bahasa Indonesia bernama IndoBERT, yang telah dilatih dengan 4 milyar data korpus atau 20 Gigabyte data teks Berbahasa Indonesia[17]. IndoNLP awalnya dimulai sebagai kolaborasi gabungan antar Universitas Multimedia Nusantara, Institut Teknologi Bandung, The Hong Kong University of Science and Technology, Universitas Indonesia, Gojek dan Prosa.ai, yang menghasilkan benchmark IndoNLU (Indonesian Natural Language Understanding)[17]. Benchmark ini mencakup kumpulan dataset dan pre-trained models berbasis Bahasa Indonesia untuk berbagai tugas pemrosesan bahasa alami, seperti analisis sentimen, klasifikasi teks, named entity recognition (NER), dan entailment. Penelitian ini bertujuan untuk meningkatkan kapabilitas model NLP dalam memahami ragam bahasa Indonesia formal maupun informal, sekaligus menyediakan tolok ukur standar untuk mengevaluasi performa model-model NLP yang dikembangkan untuk Bahasa Indonesia secara sistematis dan objektif[17].

Dalam memilih *IndoNLP*, telah dipertimbangkan beberapa algoritma lain yang umum digunakan dalam pemrosesan bahasa alami, seperti *BERT (Bidirectional Encoder Representations from Transformers)* dan *FastText*. Namun, *IndoNLP* memiliki model *IndoNLU (Indonesian Natural Language Understanding)* yang berfokus pada pemahaman bahasa alami dengan menyediakan benchmark dan model pretrained untuk tugas seperti analisis sentimen dan klasifikasi teks. IndoNLU sudah mencakup variasi model seperti IndoBERT dan IndoFastText, kedua algoritma ini sudah dilatih dan diadaptasi untuk memahami bahasa Indonesia yang beragam. *IndoBERT* merupakan Machine Learning model yang digunakan

untuk setiap model *IndoNLU*, dan model NLP yang akan digunakan untuk penelitian ini.

Berdasarkan uraian yang telah dijabarkan, akan didesain dan dirancang suatu model atau sistem dengan memanfaatkan teknologi serupa, yakni *IndoNLP*, dan secara ekstensi, *IndoNLU* dan *IndoBERT* untuk menganalisis sentimen *AI-generated artwork* di *Twitter* dengan tujuan untuk menggunakan model tersebut dalam menganalisis persepsi pengguna *Twitter* terhadap penggunaan kecerdasan buatan dalam pembuatan karya seni.

IndoNLP dipilih karena memiliki kelebihan dalam menangani keragaman gaya bahasa Indonesia formal dan informal, serta telah dioptimalkan untuk tugas-tugas pemrosesan bahasa alami yang spesifik untuk konteks Indonesia[18]. Dengan *benchmark* yang sudah disesuaikan untuk bahasa Indonesia, *IndoNLP* dapat memberikan akurasi yang lebih tinggi dalam analisis sentimen serta meminimalkan kesalahan interpretasi dalam konteks budaya dan linguistik lokal dibandingkan dengan model lain seperti GPT dan Gemini dalam melakukan sentimen analisis terhadap sentimen berbahasa Indonesia, dengan tingkat akurasi yang tinggi beserta dengan biaya implementasi (overhead cost) yang rendah[18].

1.2 Rumusan Masalah

Berdasarkan penjelasan latar belakang, maka rumusan masalah yang akan diambil pada penelitian ini adalah:

1. Bagaimana cara mengimplementasikan *IndoBERT* untuk menganalisis sentimen terhadap *AI-generated artwork* di *Twitter*?
2. Bagaimana hasil analisis sentimen dari *IndoBERT* yang bersifat positif, negatif dan netral terhadap *AI-generated artwork* di *Twitter*?

1.3 Batasan Permasalahan

Berdasarkan penjelasan latar belakang, maka rumusan masalah yang akan diambil pada penelitian ini adalah:

1. Bagaimana cara mengimplementasikan *IndoBERT* untuk menganalisis sentimen terhadap *AI-generated artwork* di *Twitter*?

2. Bagaimana hasil klasifikasi sentimen positif, negatif, dan netral yang dilakukan oleh *IndoBERT*, serta bagaimana performa model tersebut dievaluasi berdasarkan metrik seperti akurasi, F1-score, dan loss?

1.4 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah sebagai berikut:

1. Mengimplementasikan *IndoBERT* untuk menganalisis sentimen terhadap *AI-generated artwork* di *Twitter*, serta mengevaluasi performa model dengan menggunakan metrik seperti akurasi, F1-score, dan nilai loss.
2. Menginterpretasikan sentimen yang bersifat positif, negatif, dan netral terhadap *AI-generated artwork* berdasarkan hasil klasifikasi sentimen yang dilakukan oleh model *IndoBERT*.

1.5 Manfaat Penelitian

Manfaat yang diharapkan untuk dapat diperoleh dari penelitian ini adalah sebagai berikut.

1. Mendapatkan pemahaman dari analisis sentimen dari penggunaan kecerdasan buatan terhadap proses pembuatan karya seni digital menggunakan *IndoNLP*.
2. Laporan ini menjadi alat bantu bagi peneliti, *AI developer*, *Data Scientist* dan masyarakat luas dalam mengembangkan *IndoNLP* dan *NLP* untuk meningkatkan ilmu pemahaman mesin terhadap bahasa Indonesia, secara ekstensi, bahasa manusia.
3. Mendukung pembahasan penelitian lebih lanjut mengenai *NLP* tentang sentimen objek penelitian lain, seperti suatu perusahaan tertentu, atau kepuasan masyarakat terhadap pelayanan publik pemerintahan.

1.6 Sistematika Penulisan

Adapun sistematika penulisan laporan adalah sebagai berikut:

- Bab 1 PENDAHULUAN

Pada bagian ini, pendahuluan melingkupi latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan

sistematika penulisan. Penulis mengangkat fenomena kontroversial seputar penggunaan AI dalam seni digital, khususnya di media sosial Twitter, serta pentingnya menganalisis opini publik terhadapnya melalui pendekatan analisis sentimen berbasis NLP dengan model Bahasa Indonesia.

- **Bab 2 LANDASAN TEORI**

Pembahasan mengenai teori umum yang mendasari penelitian ini, mencakup konsep kecerdasan buatan (AI), platform Twitter sebagai sumber data, pengertian dan kontroversi AI-generated artwork, serta teknologi NLP (Natural Language Processing), NLU (Natural Language Understanding), dan model terkait seperti IndoBERT dan SmSA yang digunakan dalam penelitian.

- **Bab 3 METODOLOGI PENELITIAN**

Bab ini menjelaskan metodologi penelitian yang digunakan, mulai dari identifikasi masalah, pengumpulan data menggunakan Twscraper dari Twitter, proses data preprocessing (case folding, stopword removal, lemmatization, dll.), hingga penggunaan model IndoBERT dari IndoNLP untuk klasifikasi sentimen ke dalam tiga kategori: positif, netral, dan negatif.

- **Bab 4 HASIL DAN DISKUSI**

Menyajikan hasil implementasi dan evaluasi model yang digunakan. Model terbaik diperoleh dengan konfigurasi split data 40:30:30 dan epoch ke-7, dengan akurasi mencapai 75 persen. Hasil klasifikasi menunjukkan mayoritas sentimen bersifat negatif (51,8 persen), netral (47,3 persen), dan hanya 0,9 persen positif, menggambarkan kekhawatiran publik terhadap dampak AI terhadap seniman dan orisinalitas karya.

- **Bab 5 KESIMPULAN DAN SARAN**

Merangkum bahwa implementasi IndoBERT berhasil untuk tugas analisis sentimen terhadap AI-generated artwork di Twitter, dan hasil klasifikasi mengindikasikan bahwa masyarakat Indonesia secara umum bersikap skeptis terhadap fenomena tersebut. Bab ini juga memberikan beberapa saran untuk pengembang teknologi, komunitas seni, serta arah untuk penelitian selanjutnya.