

BAB I

PENDAHULUAN

1.1.Latar Belakang

Artificial Intelligence atau dapat disingkat AI sudah mulai mengubah banyak industri, salah satunya adalah industri seni. *AI-generated art* atau gambar yang diciptakan oleh AI semakin meningkat popularitasnya dalam beberapa tahun ini karena potensi kreativitasnya yang unik. Sosial media seperti Twitter sudah mejadi tempat populer untuk para seniman untuk menunjukkan hasil dari *AI-generated art* dan mendapat masukan dari public. Akan tetapi, menganalisis sentiment dari pengguna Twitter dapat menjadi sebuah tantangan karena sifat seni yang subjektif. Didalam laporan ini penulis akan mengeksplorasi bagaimana *Sentiment Analysis* digunakan untuk menganalisis sentiment dari pengguna Twitter terhadap *AI-generated art* menggunakan dua metode: metode *Neural Network* dan *Naive Bayes*.

Menurut Haiyun Peng,dkk.(2020), *Sentiment Analysis* merupakan *natural language processing* digunakan untuk menentukan sentiment dari sebuah teks. NLP merupakan alat yang kuat untuk menganalisis data sosial media, karena NLP memungkinkan penulis memahami perasaan seseorang tentang suatu topik atau kejadian. Dalam konteks Twitter, *sentiment analysis* dapat digunakan untuk menganalisa bagaimana sentimen para pengguna Twitter terhadap topik tertentu, seperti *AI-generated art*.

Klasifikasi *Neural Network* adalah tipe artificial *Neural Network* yang sering digunakan untuk proses klasifikasi. *Neural Network* terdiri dari beberapa lapisan neuron, yang masing-masing melakukan transformasi non-linear pada data masukan (F. Heimerl,dkk,2014). MLP *Neural Network* sudah terbukti sangat efektif dalam tugas sentiment analisis, karena MLP *Neural Network* dapat mempelajari hubungan kompleks diantara *input* data dan *output labels*.

Klasifikasi *Naive Bayes* adalah algoritma probabilitas yang sering digunakan untuk klasifikasi teks, termasuk analisis sentimen. Abbas, M.(2019), menjelaskan bahwa cara kerja metode ini, yaitu dengan menghitung probabilitas dari dokumen dengan kelas yang diberikan kepada dokumen atau teks tersebut dan memilih kelas dengan probabilitas tertinggi. Metode *Naive Bayes* sudah dibuktikan menjadi salah satu metode efektif dalam proses analisis sentimen, terutama ketika berhadapan dengan data teks.

Untuk melakukan *sentiment analysis* terhadap sentimen pengguna Twitter terhadap *AI-generated art*, pertama penulis akan mengambil data postingan tweets terkait dengan *AI-generated art* menggunakan Twitter API. Kemudian penulis akan melakukan tahap *preprocess data* dengan menghapus hal-hal seperti symbols, emoji, URLs, *special characters*, menghapus stop word, dan menghapus pesan spam/duplikat. Selanjutnya penulis akan melakukan stemming dan tokenization terhadap datanya untuk mengurangi dimensi dari datanya. Lalu data tersebut akan dipisah menjadi *training data* dan *test data* lalu diberikan label menggunakan Vader.

Selanjutnya, penulis akan menggunakan fungsi *TfidfVectorizer* dan *countvectorizer* dari *scikit-learn* dari *library* python untuk menkonversi data teks menjadi fitur numerik. Kedua metode *vectorizer* ini berfungsi untuk mengubah data teks menjadi fitur data matrix, yang merupakan representasi numerik dari data teks. Penulis kemudian akan menggunakan representasi numerical tersebut untuk melatih metode *MLP Neural Network* dan *Multinomial Naive Bayes*.

Untuk Metode *Neural Network*, penulis akan mendefinisikan arsitektur *Neural Network* dengan menggunakan satu *hidden layer* terdiri dari 64 neurons, menggunakan fungsi relu activation dan Adam *optimizer*. Kemudian kita akan melatih model menggunakan *training set* dan mengevaluasi performanya pada *testing set*. Selanjutnya penulis akan menghitung akurasi *precision* dan *recall* pada model menggunakan report klasifikasi dan *confusion matrix*.

Sementara itu untuk metode *Naive Bayes*, penulis akan melatih model tersebut menggunakan *training set* dan mengevaluasi performanya pada *testing set*. Selanjutnya kita akan menghitung akurasi, *precision*, dan *recall* pada model menggunakan report klasifikasi dan *confusion matrix*.

Setelah *training* dan *testing* pada model *Neural Network* dan *Naive Bayes* terhadap data set Twitter, selanjutnya akan dihitung performa kedua model dalam menghitung kemungkinan sentiment pengguna Twitter terhadap *AI-generated art*.

1.2. Rumusan Masalah

Berdasarkan latar belakang dari penelitian ini, permasalahan yang akan dibahas dalam penelitian ini adalah:

- Berapa akurasi metode klasifikasi *Multi-layer perceptron Neural Network* dalam memprediksi sentiment?
- Berapa akurasi metode klasifikasi *Multinomial Naive Bayes* dalam memprediksi sentiment?

1.3. Tujuan

Berdasarkan beberapa masalah di atas, maka tujuan dari penelitian ini adalah sebagai berikut :

- Untuk mengetahui berapa akurasi metode klasifikasi *Neural Network* dalam memprediksi sentiment dari pengguna Twitter
- Untuk mengetahui berapa akurasi metode klasifikasi *Naive Bayes* dalam memprediksi sentiment dari pengguna Twitter
- Untuk mengetahui metode mana yang memiliki akurasi yang lebih tinggi

1.4. Manfaat Penelitian

Manfaat dari penelitian ini yaitu untuk mengetahui performa dari metode klasifikasi *Neural Network* dan akurasi metode klasifikasi *Naive Bayes* dalam memprediksi sentiment dari pengguna Twitter terhadap *AI-generated art*.

1.5. Batasan Penelitian

Berikut merupakan batasan dari penelitian ini :

- Data yang digunakan adalah data dari posting user yang didapatkan dari Twitter
- Data yang digunakan menggunakan bahasa inggris
- Hasil akhir dalam penelitian ini merupakan hasil sentiment analisis dari klasifikasi *Neural Network* dan klasifikasi *Naive Bayes*
- Metode tidak dapat digunakan untuk mendeteksi kalimat yang bersifat sarkasme