

BAB I

PENDAHULUAN

1.1 Latar Belakang

Sentiment analysis atau analisis sentimen dalam bahasa Indonesia adalah sebuah teknik atau cara yang digunakan untuk mengidentifikasi bagaimana sebuah sentimen diekspresikan menggunakan teks dan bagaimana sentimen tersebut bisa dikategorikan sebagai sentimen positif maupun sentimen negative, atau netral. Hasil sistem prototipe mencapai tinggi presisi (75-95% tergantung pada data) dalam mencari sentimen pada halaman web dan artikel berita. *Sentiment analysis* atau *opinion mining* mengacu pada bidang yang luas dari pengolahan bahasa alami, komputasi linguistik dan *teks mining* yang memiliki tujuan menganalisa pendapat, sentimen, evaluasi, sikap, penilaian dan emosi seseorang apakah pembicara atau penulis berkenan dengan suatu topik, produk, layanan, organisasi, individu, ataupun kegiatan tertentu (Pangestu and Rosyda, 2022), *Sentiment analysis* yang menggunakan metode *Support Vector Machine*.

Penelitian mengenai analisis sentimen telah banyak dilakukan, terutama pada metode *Support Vector Machine*. SVM merupakan metode data mining yang masuk dalam kelas *supervised learning*. Tujuan utama dari teknik ini adalah mencari *hyperplane* terbaik dari beberapa *hyperplane* yang ada. SVM berusaha mendapatkan fungsi pemisah yang optimal untuk dapat memisahkan dua set data dari dua kelas yang berbeda. Pada teknik SVM, *hyperplane* terbaik adalah *hyperplane* yang terletak di tengah-tengah antara dua set data dari dua kelas yang berbeda. Mencari *hyperplane* terbaik adalah dengan cara memaksimalkan margin. Margin adalah jarak tegak lurus antara *hyperplane* dengan obyek terdekat. Obyek

terdekat disebut juga support vector (Styawati, Nurkholis, *et al.*, 2021). Data yang sering digunakan adalah sebuah penilaian, dan tweet di media sosial dalam hal ini di *twitter*.

Oleh karena itu, pada penelitian ini penulis akan melakukan analisis sentimen terhadap komentar netizen *twitter* yang ada di Indonesia, untuk melihat jumlah pengguna media pembelajaran AI apakah setuju dengan pembelajaran AI dan tidak setuju dengan pembelajaran AI. Data diambil dari bulan September 2022 sampai dengan Agustus 2023 berupa data *tweets* di *twitter*, penelitian ini akan menggunakan algoritma *word2vec* yang digunakan untuk memetakan setiap kata dalam teks ke dalam vector dan metode *Support Vector Machine (SVM)* yang akan digunakan sebagai proses pengklasifikasian data. Kemudian hasil tersebut akan dievaluasi untuk melihat jumlah nilai pada kelas positif dan negatif. Pada penelitian (Kusumawati, Faraby and P, 2021) yang berjudul Analisis Sentimen Komentar Beracun Pada media Sosial Menggunakan *Wor2Vec* dan *Support Machine (SVM)*, dikatakan bahwa *Word2Vec* merupakan salah satu model umum yang dapat digunakan untuk embedding kata. Di mana algoritma berbasis prediksi yang digunakan untuk mempresentasikan sebuah kata sebagai vektor dengan relasi semantik. *Word2Vec* banyak diterapkan pada *Natural Language Processing (NLP)* dan menunjukkan dampak potensial yang tinggi pada kinerja analisis sentimen. Pada penelitian yang dilakukan tersebut mendapatkan nilai *F1-Score* terbaik sebesar 73,69%.

1.2 Rumusan Masalah

Berdasarkan dari latar belakang masalah yang telah diuraikan diatas, maka rumusan masalah pada penelitian ini adalah :

1. Bagaimana klasifikasi data *tweets* diklasifikasi dengan metode *Support Vector Machine Word2vec*.
2. Bagaimana performa metode *Support Vector Machine Word2vec* terhadap *tweet Artificial Intelligence*?

1.3 Batas Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Data yang digunakan dalam penelitian ini adalah *tweet* berbahasa Indonesia.
2. Sumber Dataset berasal dari media sosial *Twitter*.
3. Data *tweets* yang akan digunakan diambil dari bulan oktober 2021 sampai dengan Juni 2023.
4. Metode yang digunakan untuk pengklasifikasian dalam penelitian ini adalah metode *Support Vector Machine (SVM)*.
5. Hasil dari pengklasifikasian sentiment pada penelitian ini menggunakan metode SVM berupa sentiment Positif dan Negatif.
6. Bahasa pemrograman yang digunakan untuk mengolah data adalah Python.

1.4 Tujuan Penelitian

Adapun tujuan yang ingin dicapai oleh penulis dari penelitian ini adalah :

1. Melakukan analisa terhadap data *Artificial Intelligence*.
2. Mengklasifikasikan data *Artificial Intelligence* dan menentukan kelas data *tweets* dari pengguna *Twitter* menggunakan algoritma *SVM Word2vec*.
3. Melakukan pengujian metode *SVM Word2vec* menggunakan data dari opini Masyarakat terhadap *Artificial Intelligence*.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah:

1. Diharapkan dapat memberikan manfaat baik kepada penulis maupun pembaca tentang gambaran sentiment para pengguna *Twitter* terhadap *Artificial Intelligence*.
2. Untuk Mengetahui performa metode *Support Vector Machine Word2vec* dalam melakukan proses klasifikasi data terhadap *Artificial Intelligence*.