

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Analisis sentimen adalah salah satu area penelitian yang sangat aktif dilakukan [1]. Analisis sentimen ditujukan untuk mendeteksi, menganalisa, dan mengevaluasi pikiran manusia terhadap suatu hal tertentu berdasarkan tulisan, ekspresi wajah, cara bicara, musik, gerakan, dan lain-lain. [2]. Berdasarkan pengertian dari kamus Merriam-Webster, sentimen adalah sebuah sikap, pemikiran, atau penilaian yang didorong oleh sebuah perasaan. Analisis sentimen terbagi menjadi 2 (dua) yaitu *opinion mining* dan *emotion mining* [2].

Deteksi emosi secara khusus telah menjadi fokus penelitian dalam bidang *emotion mining* [3]. Emotion mining adalah penelitian terhadap suatu emosi (contoh, senang, sedih, marah) yang tergambar dari suatu input. Ada banyak penelitian yang telah dilakukan di bidang deteksi emosi dengan berbagai jenis *input*, dimana salah satunya berasal dari teks [4]. Deteksi emosi dari teks adalah tugas yang relatif lebih kompleks jika dibandingkan dengan deteksi sentimen [4].

Saat ini, sistem deteksi emosi berdasarkan teks bahasa Indonesia mulai banyak dilakukan [5]–[8], namun terdapat kendala dalam pengembangan model deteksi emosi teks bahasa Indonesia yaitu bahasa Indonesia termasuk dalam golongan '*under-resource language*' [1]. Hal ini dikarenakan kurang tersedianya korpus atau sumber data teks bahasa Indonesia yang dapat digunakan untuk deteksi emosi [3]. Beberapa peneliti telah menyediakan dataset/korpus yang digunakan

oleh mereka dalam penelitiannya, namun jumlah data yang tersedia tetap saja tidak cukup [4]. Selain itu, belum ada dataset yang *standard* yang dapat digunakan untuk model deteksi emosi [4]. Oleh karena itu, para peneliti terikat untuk menggunakan dataset pada model deteksi emosi yang sudah ada atau membuat dataset baru dan memberikan label pada data secara manual dimana hal ini menghabiskan banyak waktu untuk dilakukan [4].

Saat ini terdapat beberapa cara atau teknik yang digunakan untuk mengembangkan model deteksi emosi berdasarkan teks bahasa Indonesia [6]–[8]. Pengembangan model deteksi emosi menggunakan komentar Youtube sebagai sumber data dan *word embedding* dengan metode CNN menghasilkan akurasi 76.2% [6]. Selain itu pengembangan model deteksi emosi dengan metode LDA (*Latent Dirichlet Allocation*) dan konversi simbol ekspresi (*emoji*) pada *tweet* bahasa Indonesia memberikan akurasi terbaik sebesar 50.8% [7]. Pengembangan model deteksi emosi menggunakan algoritma Naïve Bayes dan dataset yang dikembangkan oleh Julian Bata dengan menggunakan *hashtag* (#) pada *tweet* sebagai anotasi menghasilkan akurasi sebesar 82% [3]. Sebuah penelitian juga dilakukan untuk membangun dataset deteksi emosi teks bahasa Indonesia yang dikumpulkan dari *tweet* bahasa Indonesia, sebagai tambahan penelitian tersebut juga mengusulkan beberapa teknik yang dapat digunakan pada dataset tersebut guna melakukan klasifikasi emosi dimana hasilnya menunjukkan bahwa penggunaan algoritma *Logistic Regression* dan kombinasi *features* pada dataset tersebut memberikan akurasi terbaik yaitu 69.73% [8].

Oleh karena itu, berdasarkan paparan di atas, penelitian ini akan berfokus untuk membangun model deteksi emosi menggunakan dataset yang digunakan pada [8] dengan menjadikan akurasi terbaik pada penelitian tersebut sebagai *baseline* dalam penelitian ini.

1.2. Rumusan Masalah

Berdasarkan latar belakang masalah di atas, masalah yang akan dibahas/diteliti dalam penelitian ini adalah sebagai berikut:

- a. Bagaimana model deteksi emosi pada teks bahasa Indonesia dibangun?
- b. Bagaimana hasil akurasi model deteksi emosi pada teks bahasa Indonesia yang telah dibangun?

1.3. Batasan Masalah

Berikut adalah batasan masalah yang ditetapkan dalam penelitian ini:

- a. Dataset yang digunakan untuk pengembangan model adalah data teks bahasa Indonesia yang digunakan pada penelitian sebelumnya [8].
- b. Emosi yang digunakan dalam penelitian ini hanya ada 5 jenis emosi yaitu *happy, love, fear, sadness, dan anger*.
- c. *Baseline* pada penelitian ini adalah akurasi terbaik dari penelitian sebelumnya yaitu 69.73% [8].

1.4. Tujuan dan Mafaat Penelitian

1.4.1. Tujuan Penelitian

Tujuan dari penelitian ini adalah:

- a. Membangun model deteksi emosi pada teks bahasa Indonesia.
- b. Mengetahui hasil akurasi model deteksi emosi pada teks bahasa Indonesia yang telah dibangun.

1.4.2. Manfaat Penelitian

Manfaat dari penelitian ini adalah:

- a. Sebagai referensi bagi penelitian yang ditujukan untuk pengembangan model deteksi emosi pada teks bahasa Indonesia.
- b. Menambah wawasan mengenai penggunaan algoritma *machine learning* dalam membangun model deteksi emosi.