

BAB I

PENDAHULUAN

1.1 Latar Belakang

Sebuah teks dapat berisi suatu informasi tentang perilaku manusia terhadap emosi. Jenis emosi seperti normal, senang, sedih, marah dan jijik telah di kenal sejak lama dan menjadi peranan dalam komunikasi antar manusia di kehidupan sehari hari dan menjadi aspek penting dari perilaku manusia pada umumnya. Akan tetapi penerapan emosi belum banyak di gunakan untuk interaksi antara manusia dan komputer. Teks emosi tidak hanya ditunjukkan melalui lisan, tetapi juga dapat melalui tulisan. Emosi seseorang lebih mudah ditebak karena tergambar melalui ekspresi wajah dan nada suara seseorang pada saat berbicara. (Mursyidah and Hidayat, 2017)

Seiring dengan perkembangan teknologi yang semakin canggih, dilakukan penelitian di bidang emosi antara manusia dan komputer. Penelitian yang telah di lakukan sebagian besar masih menggunakan bahasa inggris dan bahasa Indonesia, sedangkan untuk teks bahasa batak belum dilakukan. Oleh karena itu penulis akan membuat penelitian klasifikasi teks emosi menggunakan bahasa batak.

Untuk mengklasifikasikan emosi yang terdapat pada teks tersebut menggunakan metode TF-IDF dan *Support Vector Machine*. Dengan metode TF-IDF akan dilakukan preprocessing dan pembobotan kata . kemudian hasilnya diklasifikasikan dengan algoritma *support vector machine* tersebut dan dengan metode ini juga diharapkan bisa mendapat hasil yang tingkat keakuratnya tinggi. (Abirawa & Osmond, 2018)

Penelitian dengan penerapan metode *Term Frekuensi/ Invers Dokumen Frekuensi* sudah pernah dilakukan dengan menggunakan bahasa yang berbeda. Seperti pada penelitian yang dilakukan oleh (Mursyidah and Hidayat, 2017) mengenai Klasifikasi Teks Emosi Bahasa Aceh Menggunakan Metode *Term Frekuensi/ Invers Dokument Frekuensi*. Data yang digunakan adalah bahasa Aceh, *Term Frekuensi/ Invers Dokument Frekuensi* (TF-IDF) adalah metode yang

digunakan untuk menghitung bobot setiap kata yang paling umum digunakan pada *information retrieval*. Dalam proses menghitung bobot setiap kata teks emosi menggunakan bahasa aceh disimpan dalam format *.txt* di dalam suatu folder tersendiri berdasarkan kelasnya masing-masing. Setelah data teks disimpan tahap selanjutnya *Indexing* untuk memberikan nomor pada setiap kata, tahap *tokenizing* untuk memecah kata berdasarkan spasi, tahap *filtering* digunakan untuk membuang *stopword* yaitu kata penghubung yang akan dihapus karena termasuk kata yang tidak mempunyai arti. Penelitian sistem aplikasi ini menghasilkan akurasi kebenaran prediksi sebesar 80%.

Penelitian yang dilakukan oleh (Abirawa & Osmond, 2018) mengenai Klasifikasi Emosi Pada lirik Lagu Menggunakan Metode *Support Vector Machine*. Terdapat dua kelompok penelitian yang mempelajari hubungan antara musik dengan emosi, yaitu Psikologi Musik, mereka yang mempelajari hubungan antara isyarat akustik (*beat*, tempo tingkat suara, dll.) dengan berbagai ekspresi emosi (marah, sedih, senang, dll.) dan sebagian besar mereka yang mengembangkan model emosional. Penelitian bidang komputer, yaitu mereka yang mengembangkan algoritma untuk mendeteksi emosi terhadap musik secara otomatis. Mereka mencoba menggunakan emosi selain untuk mendapatkan *metadata* konvensional seperti jenis musik, musik mood, yang juga sebagai penelitian dalam bidang *Music Information Retrieval*. Pada penelitian ini untuk mengklasifikasikan emosi yang terdapat pada lirik lagu tersebut menggunakan metode *Support Vector Machine* dan pembobotan menggunakan TF-IDF. Dengan metode *Support Vector Machine* tersebut dapat menentukan klasifikasi emosi apakah sesuai pada lirik lagu tersebut dan dengan metode ini juga diharapkan bisa mendapat hasil yang tingkat keakuratannya tinggi.

Penelitian yang dilakukan oleh (Fazlur Rahman, 2018) mengenai Klasifikasi Emosi Untuk Teks Berbahasa Indonesia Pada Pengguna Twitter Mengenai Presiden Joko Widodo. Penelitian ini ditujukan untuk dapat mengetahui opini dan emosi dari masyarakat. Tanggapan publik mengenai Presiden Joko Widodo didapat dari *Application Programming Interface (API)*. Sebelum dilakukan klasifikasi teks, akan dilakukan praproses teks. Praproses teks yang digunakan

adalah *case folding*, *stopwords*, *tokenizing*. Sedangkan pada analisis klasifikasi data teks tersebut digunakan metode *Support Vector Machine* (SVM) dan *K-Nearest Neighbor* (KNN). Klasifikasi menggunakan SVM menghasilkan akurasi prediksi sebesar 95,2% sedangkan KNN menghasilkan akurasi 87,2%.

1.2 Rumusan Masalah

Rumusan masalah yang terdapat di dalam tugas ini yaitu sulitnya mengklasifikasi emosi sebuah teks. Penelitian-penelitian sebelumnya yang melakukan klasifikasi teks emosi berbahasa inggris, indonesia aceh. Pada penelitian ini adalah penyempurnaan pada penelitian-penelitian sebelumnya dengan teks berbahasa batak toba dan memiliki tahapan *preprocessing* dengan proses *tokenizing* dan *filtering* kemudian tahapan klasifikasi menggunakan metode *Support Vector Machine* (SVM).

1.3 Batasan Masalah

Adapun batasan masalah dari penelitian ini adalah sebagai berikut:

1. Emosi yang diklasifikasikan hanya 5 emosi secara umum malu, jijik, marah, senang, sedih.
2. Dataset yang digunakan adalah teks emosi berbahasa batak toba.
3. Metode yang digunakan dalam klasifikasi adalah *Support Vector Machine*

1.4 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah:

1. Mendapatkan suatu klasifikasi emosi yang ada dalam teks
2. Menganalisa nilai persentase keberhasilan dari metode *Support Vector Machine* yang digunakan
3. Mendapatkan kata-kata yang sering muncul pada setiap kategori emosi

1.5 Sistematika Penulisan

Adapun sistematika penulisan dalam penelitian ini adalah sebagai berikut:

BAB I PENDAHULUAN

Berisi tentang latar belakang, perumusan masalah, batasan masalah, tujuan penelitian dan sistematika penulisan.

BAB II LANDASAN TEORI

Bab ini berisikan teori yang berupa pengertian dan defenisi yang diambil dari kutipan buku yang berkaitan dengan penyusunan laporan skripsi serta beberapa literatur riview yang berhubungan dengan penelitian.

BAB III METODOLOGI PENELITIAN

Bab ini berisikan tahapan penelitian yang akan dilakukan seperti studi literature, identifikasi masalah, pengumpulan data, perancangan sistem, pengujian dan tahapan evaluasi.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menjelaskan analisa sistem yang diusulkan dengan menggunakan flowchart dari sistem yang diimplementasikan, serta pembahasan secara detail dari bab sebelumnya, dijabarkan secara satu persatu dengan menerapkan konsep sesudah adanya sistem yang diusulkan.

BAB V PENUTUP

Bab ini berisi kesimpulan dan saran yang berkaitan dengan analisa dan optimalisasi sistem berdasarkan yang telah diuraikan pada bab-bab sebelumnya.

DAFTAR PUSTAKA

LAMPIRAN