BAB 1
PENDAHULUAN

1.1 Latar Belakang Masalah

Teks merupakan salah satu media yang digunakan untuk berkomunikasi dan menyampaikan informasi [1]. Tidak hanya memuat informasi, teks juga dapat mengekspresikan emosi [2]. Lirik lagu merupakan salah satu bentuk teks yang dapat digunakan sebagai objek dalam penelitian klasifikasi emosi. Dalam penentuan emosi, lirik lagu merupakan elemen yang memiliki makna paling kuat dalam menggambarkan emosi [3].

Penelitian terhadap klasifikasi emosi pada lirik lagu Bahasa Indonesia pernah dilakukan sebelumnya oleh Lailatus Sofiyana menggunakan metode *K-Nearest Neighbor* (K-NN) dengan nilai akurasi mencapai 60% [4]. Penelitian tersebut belum menghasilkan tingkat akurasi yang maksimal karena performansi metode klasifikasi yang digunakan kurang baik dalam mengklasifikasikan teks. Hal ini ditunjukkan pada penelitian klasifikasi teks Bahasa Indonesia yang dilakukan oleh Fatimah Wulandini bahwa metode K-NN memberikan hasil akurasi paling rendah dibandingkan metode klasifikasi *Support Vector Machine*, *Naïve Bayes Classifier*, dan *C4.5 Decision Tree* dengan akurasi sebesar 49,17% [5].

Support Vector Machine (SVM) merupakan metode klasifikasi yang berakar dari teori pembelajaran statistik yang hasilnya sangat menjanjikan untuk memberikan hasil yang lebih baik dari metode yang lain. Selain itu, SVM juga bekerja dengan baik pada set data dengan dimensi yang tinggi [6], seperti teks. Hal ini dibuktikan dengan penelitian Fatimah Wulandini yang berjudul "*Text Classification Using Support Vector Machine for Webmining Based Spatio Temporal Analysis of the Spread of Tropical Diseases"* bahwa metode SVM menunjukkan hasil paling baik pada kasus klasifikasi informasi teks Bahasa Indonesia dibandingkan dengan metode *Naive Bayes classifier*, *k-Nearest Neighbor classifier*, dan C4.5 *Decision Tree* dengan akurasi sebesar 92,5%. Penelitian tersebut berhasil mengklasifikasikan 111 dari 120 artikel Bahasa Indonesia secara

benar sesuai dengan kelas asli ke dalam 6 kelas kategori, yaitu *economy, defense & security, education, health, sports*, dan *politics* [5]. Berdasarkan hasil dari penelitian tersebut membuktikan bahwa metode SVM dapat memecahkan masalah klasifikasi *multi class* dengan baik.

Berdasarkan uraian diatas, maka pada penelitian ini digunakan metode *Multi Class Support Vector Machine* untuk mengklasifikasikan emosi pada lirik lagu Bahasa Indonesia.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, maka rumusan masalah dalam penelitian ini adalah bagaimana cara mengimplementasikan metode *Multi Class Support Vector Machine* untuk klasifikasi emosi pada lirik lagu Bahasa Indonesia.

1.3 Maksud dan Tujuan

Maksud dari penelitian ini adalah mengimplementasikan metode *Multi Class Support Vector Machine* untuk klasifikasi emosi pada lirik lagu Bahasa Indonesia. Sedangkan tujuan yang ingin dicapai dari penelitian ini adalah mengetahui besar akurasi yang didapatkan dengan mengimplementasikan metode *Multi Class Support Vector Machine* untuk klasifikasi emosi pada lirik lagu Bahasa Indonesia.

1.4 Batasan Masalah

Agar penelitian yang dilakukan lebih terarah dan mencapai sasaran yang ditentukan, maka diperlukan sebuah pembatasan masalah atau ruang lingkup kajian, yaitu sebagai berikut:

1. Data latih yang digunakan adalah data dari *International Survey On Emotion Antecedents And Reaction* (ISEAR) berjumlah 1000 data pernyataan dengan format (.txt) [7].
2. Data uji yang digunakan adalah lirik lagu Bahasa Indonesia berjumlah 30 data dengan format (.txt) [4] [8] [9].
3. Kategori emosi yang digunakan adalah senang, takut, marah, sedih, dan bersalah, disesuaikan dengan penelitian sebelumnya [4].

4. Tahapan *preprocessing* meliputi *case folding, convert negation, tokenizing, stopword removal*, dan *stemming* [10] [11].
5. Algoritma *stemming* yang digunakan adalah algoritma Nazief & Adriani [12].
6. Metode pembobotan yang digunakan adalah TF-IDF [10].
7. Metode *Multi Class Support Vector Machine* yang digunakan adalah *one against all* [6].
8. Fungsi kernel yang digunakan adalah linear [6].

1.5 Metodologi Penelitian

Metodologi penelitian merupakan cara utama yang digunakan peneliti untuk mencapai tujuan dan menentukan jawaban atas masalah yang diajukan [13]. Metode penelitian yang digunakan dalam penelitian ini adalah metode eksperimen. Metode eksperimen didefinisikan sebagai metode sistematis guna membangun hubungan yang mengandung fenomena sebab akibat. Metode eksperimen merupakan metode inti dari model penelitian yang menggunakan pendekatan kuantitatif [14]. Variabel bebas yang diteliti adalah metode klasifikasi *Multi Class Support Vector Machine* yang digunakan untuk klasifikasi emosi pada lirik lagu Bahasa Indonesia, sedangkan variabel terikatnya adalah nilai akurasi.

Penelitian ini memiliki 3 metode dalam pelaksanaannya, yaitu metode pengumpulan data, metode pembangunan perangkat lunak, dan metode klasifikasi. Berikut ini merupakan penjelasan dari masing-masing metode:

1.5.1 Metode Pengumpulan Data

Metode pengumpulan data yang digunakan dalam penelitian ini adalah sebagai berikut:

a. Studi Literatur

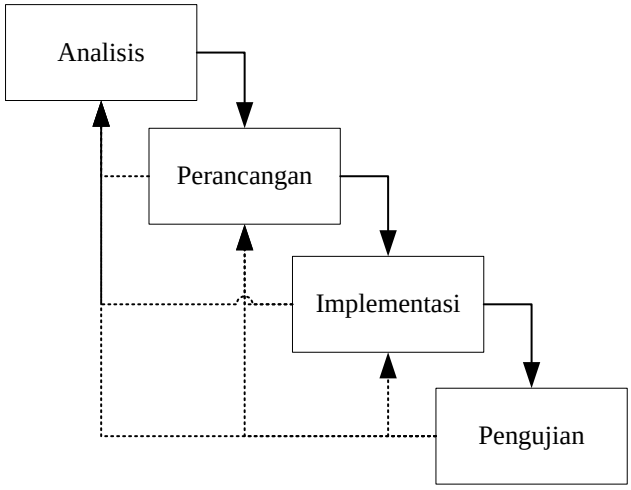
Studi literatur yang dilakukan adalah pengumpulan data dengan cara mengumpulkan jurnal, *paper*, buku, serta bacaan-bacaan yang dapat mendukung sebagai referensi pada penelitian ini. Studi literatur yang dibutuhkan adalah mengenai metode *Multi Class Support Vector Machine*.

b. Observasi

Observasi yang dilakukan adalah pengumpulan data dari *International Survey On Emotion Antecedents And Reaction* (ISEAR) yang digunakan sebagai data latih dan lirik lagu Bahasa Indonesia sebagai data uji.

1.5.2 Metode Pembangunan Perangkat Lunak

Metode pembangunan perangkat lunak merupakan langkah-langkah yang akan dilakukan dalam membangun simulasi untuk mengklasifikasikan emosi pada lirik lagu Bahasa Indonesia. Berikut ini adalah tahapan metode pembangunan perangkat lunak yang akan dilakukan dalam penelitian ini, dapat dilihat pada Gambar 1.1



Gambar 1.1 Metode Pembangunan Perangkat Lunak

Berikut merupakan penjelasan masing-masing tahapan dari Gambar 1.1:

1. Analisis

Dalam tahapan ini dilakukan untuk mempelajari konsep dan analisis mengenai klasifikasi emosi menggunakan metode *Multi Class Support Vector Machine*.

2. Perancangan

Pada tahap ini dilakukan perancangan dari simulasi klasifikasi emosi pada lirik lagu Bahasa Indonesia yang akan dibuat dengan analisis yang telah dilakukan pada tahap sebelumnya.

3. Implementasi

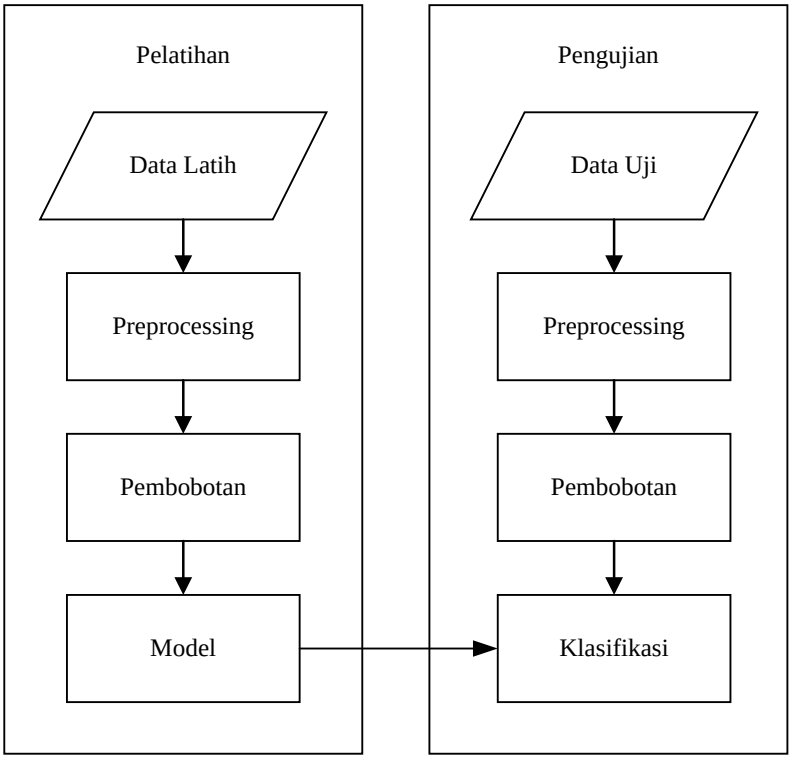
Pada tahap ini akan dilakukan proses implementasi metode *Multi Class Support Vector Machine* ke dalam simulasi yang akan dibangun sesuai dengan perancangan yang dibuat pada tahap sebelumnya.

4. Pengujian

Pada tahap ini akan dilakukan pengujian terhadap simulasi yang telah dibangun, apakah simulasi yang dibangun sudah sesuai dengan tujuan dari penelitian.

1.5.3 Metode Klasifikasi

Berikut adalah gambaran metode klasifikasi yang dilakukan dalam penelitian ini dapat dilihat pada Gambar 1.2



Gambar 1.2 Metode Klasifikasi

Berikut penjelasan dari masing-masing tahapan Gambar 1.2:

1. *Preprocessing*

Data latih dan data uji yang akan digunakan dilakukan proses *preprocessing* untuk mengubah bentuk data yang belum terstruktur menjadi data yang terstruktur. Proses *preprocessing* ini terdiri dari 5 tahap yaitu *case folding, convert negation, tokenizing, stopword removal* dan *stemming*.

2. Pembobotan

Data latih dan data uji yang telah di*preprocessing* akan dilakukan pembobotan menggunakan metode pembobotan TF-IDF.

3. Model

Data latih yang telah melalui proses *preprocessing* dan pembobotan akan dijadikan model untuk acuan dalam memprediksikan kelas teks.

4. Klasifikasi

Data uji yang telah melalui proses *preprocessing* dan pembobotan akan dilakukan klasifikasi menggunakan model yang dihasilkan pada tahap pelatihan. Metode klasifikasi yang digunakan pada penelitian ini adalah *Multi Class Support Vector Machine*.

1.6 Sistematika Penulisan

Sistematika penulisan penelitian ini disusun untuk memberikan gambaran umum mengenai penelitian yang dikerjakan. Sistematika penulisan penelitian adalah sebagai berikut:

BAB 1 PENDAHULUAN

Bab ini menjelaskan tentang gambaran umum mengenai isi laporan skripsi yang berisi penjelasan meliputi latar belakang masalah, rumusan masalah, maksud dan tujuan, batasan masalah, metodologi penelitian, serta sistematika penulisan dalam penelitian tentang implementasi metode *Multi Class Support Vector Machine* untuk klasifikasi emosi pada lirik lagu Bahasa Indonesia.

BAB 2 LANDASAN TEORI

Bab ini menjelaskan tentang berbagai konsep dan teori-teori para ahli dari literatur pendukung untuk menampilkan kategori emosi dari lirik lagu Bahasa Indonesia dengan mengimplementasikan metode *Multi Class Support Vector Machine*.

BAB 3 ANALISIS DAN PERANCANGAN

Bab ini berisi penjelasan mengenai analisis masalah dari objek penelitian yang dilakukan untuk mengetahui masalah apa yang timbul dan mencoba memecahkan masalah tersebut, analisis proses, analisis kebutuhan data masukan, analisis proses *prepocessing*, analisis pembobotan TF-IDF, analisis klasifikasi emosi dengan metode *Multi Class Support Vector Machine*, dan pengukuran kinerja klasifikasi dari penggunaan metode *Multi Class Support Vector Machine* untuk klasifikasi emosi pada lirik lagu Bahasa Indonesia. Selain itu dijelaskan juga analisis kebutuhan non fungsional, analisis kebutuhan fungsional, dan perancangan struktur menu, perancangan antarmuka, perancangan pesan, dan jaringan semantik dari perangkat lunak yang akan dibangun.

BAB 4 IMPLEMENTASI DAN PENGUJIAN

Bab ini berisi tentang implementasi dari analisis dan perancangan yang telah dilakukan pada bab sebelumnya. Implementasi yang dilakukan terdiri dari implementasi sistem berupa implementasi perangkat keras, implementasi perangkat lunak, dan implementasi antarmuka. Selain itu dilakukan pengujian terhadap perangkat lunak yang dibangun melalui rencana pengujian, skenario pengujian, serta mengukur kinerja klasifikasi emosi yang dihasilkan. Informasi yang ditampilkan dari implementasi dan pengujian yang dilakukan berupa sebuah kategori emosi dari lirik lagu.

BAB 5 KESIMPULAN DAN SARAN

Bab ini berisi penjelasan mengenai kesimpulan yang didapatkan dari pembahasan pada bab-bab sebelumnya, serta saran yang dapat diberikan oleh peneliti untuk penelitian atau pengembangan selanjutnya. Pada bagian kesimpulan berisi tentang apakah dengan penerapan metode *Multi Class Support Vector Machine* dapat mengklasifikasikan emosi pada lirik lagu Bahasa Indonesia atau tidak. Sedangkan bagian saran berisi tentang hal-hal yang belum bisa terselesaikan pada penelitian yang telah dilakukan, serta hal-hal yang bisa dilakukan untuk pengembangan selanjutnya.