

1. PENDAHULUAN

1.1. Latar Belakang

Klasifikasi teks saat ini menjadi topik yang sangat populer karena kebutuhan untuk suatu dokumen elektronik yang semakin meningkat [1]. Saat ini, klasifikasi teks sudah berhasil diterapkan pada berbagai domain seperti deteksi emosi [2], penyaringan spam SMS [3], klasifikasi situs web [4], dan analisis sentiment [5]. Emosi merupakan perasaan seseorang yang dapat mempengaruhi tingkah laku. Emosi dapat dipicu oleh lingkungan sekitar. Salah satu lingkungan yang dapat memicu emosi adalah dari lagu yang didengarkan. Para psikolog ahli menyatakan bahwa bagian otak manusia yang digunakan untuk merasakan musik berdekatan dengan bagian otak yang berhubungan dengan ekspresi emosional manusia, sehingga terdapat keterkaitan antara musik dan emosi [6]. Dalam suatu penelitian juga dinyatakan bahwa setiap musik memiliki emosi yang terpancar, maka dari itu banyak penelitian yang telah dilakukan terkait pengenalan emosi musik. Dari sudut pandang psikolog musik, emosi adalah suatu elemen yang paling penting untuk penggolongan musik itu sendiri [7].

Dalam penelitian tentang keterkaitan musik dengan emosi terdapat beberapa kelompok yang melakukan penelitian tersebut, salah satunya yaitu peneliti bidang komputer yang membuat algoritma untuk mendeteksi emosi secara otomatis [6] [8], yang salah satunya dilakukan dengan klasifikasi teks. Pada umumnya klasifikasi teks terdiri dari beberapa proses yaitu preprocessing, ekstraksi fitur, seleksi fitur, dan klasifikasi. Pada tahap preprocessing sendiri biasanya terdapat proses penyaringan, tokenisasi, dan pembuatan akar kata (normalisasi). Beberapa penelitian [9]-[10] menyatakan bahwa ekstraksi fitur, seleksi fitur, dan metode klasifikasi memiliki dampak yang besar terhadap suksesnya proses klasifikasi teks, selain itu tahap preprocessing juga mempunyai dampak yang nyata tergantung dari kombinasi domainnya. Jika kombinasi domain baik maka akan meningkatkan akurasi klasifikasi sedangkan kombinasi yang tidak tepat akan menurunkan akurasi [1]. Pada preprocessing terdapat beberapa teknik dalam pembuatan akar kata, Stemming dan Lemmatization yang sejak tahun 1960 sudah dipelajari, dan algoritmanya sudah mulai dikembangkan [11]. Lemmatization sendiri bertujuan untuk mengubah kata-kata yang sudah diubah ke bentuk kamusnya. Dari beberapa jurnal yang terkait, hanya ada metode stemming untuk Bahasa Jawa yang sudah dikembangkan [11]. Stemming sendiri bertujuan untuk mengurangi jumlah variasi dari suatu Bahasa ke dalam bentuk standar, dan representasi kanonik (stem) [13]. Pada penelitian tentang implementasi stemmer Bahasa Jawa, kata jadian yang sebelumnya telah diberi awalan, sisipan, maupun akhiran akan dikembalikan ke bentuk dasarnya. Penelitian tersebut dilakukan untuk menemukan informasi terhadap dokumen teks berbahasa Jawa yang secara struktur kata dan data akan berbeda dengan lirik lagu berbahasa Jawa sehingga ada terdapat ketidakcocokan penggunaan stemmer tersebut untuk proses klasifikasi lirik lagu. Pada penelitian tersebut stemmer Bahasa Jawa belum dapat mengembalikan kata dasar yang tepat sehingga pada beberapa kata imbuhan masih belum dapat dikembalikan kata dasarnya dikarenakan struktur kata Bahasa Jawa yang cukup rumit. Maka dari itu penelitian ini dilakukan untuk menutupi hal tersebut, selain itu dengan dirancangnya lemmatizer Bahasa Jawa ini diharapkan dapat membantu penelitian selanjutnya terkait klasifikasi teks berbahasa Jawa.

1.2. Perumusan Masalah

Rumusan masalah dalam tugas akhir ini adalah sebagai berikut :

- a) Apakah *lemmatizer* bahasa jawa berpengaruh pada proses klasifikasi lirik lagu berbahasa jawa?
- b) Apakah pengaruh *lemmatizer* bahasa jawa lebih baik dari tanpa menggunakan *lemmatizer* dalam proses klasifikasi lirik lagu berbahasa jawa?

1.3. Tujuan

Tujuan dari tugas akhir ini adalah sebagai berikut :

- a) Mengetahui pengaruh *lemmatizer* bahasa jawa terhadap proses klasifikasi lirik lagu berbahasa jawa.
- b) Mengetahui pengaruh mana yang lebih baik antara *lemmatizer* bahasa jawa dengan tanpa *lemmatization* dalam proses klasifikasi lirik lagu berbahasa jawa.

1.4. Rencana Kegiatan

Rencana kegiatan yang akan dilakukan adalah sebagai berikut :

- Studi Literatur
Mempelajari konsep – konsep yang akan digunakan pada proses klasifikasi teks dan *preprocessing*.
- Pengumpulan Data
Mengumpulkan data – data terkait lirik lagu berbahasa jawa.
- Pembangunan Sistem
Merancang dan membangun sistem.
- Analisis Hasil Sistem
Melakukan percobaan dan analisis dari hasil percobaan.
- Penulisan Laporan
Menuliskan hasil dan kesimpulan keseluruhan yang didapat dari hasil percobaan yang dilakukan.

