

BAB I

PENDAHULUAN

1.1 Latar Belakang

Twitter adalah salah satu media sosial yang paling banyak digunakan di seluruh dunia, yang memungkinkan pengguna untuk berbagi informasi dan opini dengan sesamanya secara realtime. Informasi yang dibagikan pada twitter biasanya disebut dengan kicauan (Tweet) yang terdiri dari 140 karakter. Pengguna Twitter menuliskan pendapat dan opini tentang berbagai layanan di Twitter, karena pengguna twitter bisa mengekspresikan pendapat dan opini tentang apapun termasuk *genre musik*, pengguna memiliki beragam pemikiran atas opininya tersebut, kicauan atau (Tweet) akan bertambah dan semakin banyak setiap saat dan penyebaran opini tentang genre musik pun tersebar dengan cepat. Salah satu contoh media sosial tempat bertukar opini adalah Twitter. Dari twitter dapat dilakukan Opinion Mining atau biasa disebut Analisis Sentimen. Untuk melakukan Opinion Mining di Twitter peneliti akan mengambil data yang dengan memanfaatkan API Twitter.

Analisis Sentimen merupakan suatu proses pengolahan data berupa teks yang bertujuan untuk mengklasifikasi sentimen positif dan negatif. Dalam proses pengolahan data sentimen akan melewati beberapa tahapan yaitu Pengumpulan Data, Cleaning Data, Tokenization, Stopword Removal, Normalisasi Bahasa, Stemming, Pembobotan kemudian setelah itu di klasifikasi menggunakan algoritma Support Vector Machine. SVM merupakan salah satu metode klasifikasi dengan

menggunakan metode machine learning (supervised learning) yang memprediksi kelas berdasarkan pola dari hasil proses training yang diciptakan oleh Vladimir Vapnik. Klasifikasi dilakukan dengan garis pembatas (hyperlane) yang memisahkan antara kelas opini positif dan opini negatif [1].

Berdasarkan latar belakang masalah diatas, maka Metode Support Vector Machine (SVM) ini bisa memberikan nilai potensial yang dapat diimplementasikan untuk analisis sentimen yang bertujuan guna meningkatkan strategi dalam pemasaran, peningkatan minat pada suatu genre dan semakin banyak yang mengetahui genre musik baru. Diharapkan adanya sistem ini dapat meningkatkan penikmat musik dan tidak membedakan-bedakan setiap genre musik.

1.2 Rumusan Penelitian

Penulis akan mengemukakan beberapa rumusan penelitian yang terkait dengan latar belakang pada point 1.1, yaitu sebagai berikut:

1. Bagaimana menemukan sentimen yang diberikan pengguna twitter terhadap genre musik menggunakan Metode Support Vector Machine.
2. Mencari nilai akurasi dengan metode *Confusion Matrix* untuk hasil Support Vector Machine pada klasifikasi sentimen terhadap genre musik.

1.3 Batasan Penelitian

Dalam penelitian ini memiliki batasan penelitian sebagai berikut :

1. Penelitian hanya dilakukan terhadap data dari twitter dengan keyword 'Musik', 'kpop', 'pop'.
2. Bahasa yang diambil dari twitter adalah Bahasa Indonesia

3. Analisis Sentimen dalam penelitian ini menggunakan data tweet dari twitter yang di unduh selama dua kali dalam rentang waktu seminggu.
4. Analisis Sentimen dengan metode SVM menggunakan tools yang tersedia di library Python.
5. Diklasifikasikan ke dalam 2 kelas yaitu Positif dan Negatif
6. Bahasa Pemrograman yang digunakan adalah Python versi 3.7
7. Analisis Sentimen untuk data latih diunduh dari twitter dan mendapatkan *.json* serta data uji yang diunduh secara realtime dari twitter.
8. Dalam penelitian ini belum meneliti pengukuran atau presentase berapa orang yang menyukai genre musik tertentu.
9. Dalam penelitian ini belum meneliti genre musik apa saja yang diminati tahun ini.

1.4 Tujuan Penelitian

Tujuan dari penelitian yang hendak dicapai adalah :

1. Menerapkan metode Support Vector Machine pada Analisis Sentimen terhadap Genre Musik di twitter.
2. Menemukan Sentimen pengguna twitter terhadap *genre musik* dari komentar twitter menggunakan metode Support Vector Machine.
3. Untuk mengetahui nilai presentase tingkat akurasi metode Support Vector Machine pada penelitian ini dengan menggunakan metode Confusion Matrix.

1.5 Metodologi Penelitian

Dalam penelitian ini akan menjelaskan tahapan Analisis Sentimen sebagai berikut.

1.5.1 Metode Pengumpulan Data

Metode pengumpulan data adalah metode yang digunakan untuk memngumpulkan data yang diperlukan dalam penelitian. Beberapa cara untuk mengumpulan data-data yang diperlukan penulis diantaranya :

1. Studi Pustaka

Mencari jurnal atau literature yang berkaitan dengan Analisis Sentimen, Algoritma Support Vector Machine, Confusion Matrix.

2. Kuesioner

Memberi beberapa pertanyaan kepada pengguna twitter secara acak tentang Genre Musik dengan menggunakan Kuesioner melalui *Google Form* yang disediakan oleh *google* untuk membuat Kuesioner.

1.5.2 Tahapan Penelitian

Metode pengembangan sistem disusun berdasarkan data-data yang sudah diperoleh dari metode pengumpulan data. Ada beberapa tahapan penelitian yaitu :

a. Analisis Sistem

Pada tahapan ini penulis melakukan pengkatagorian variable mana saja yang menjadi kebutuhan fungsional dan non fungsional serta kebutuhan antarmuka dengan menghasilkan analisis.

b. Perancangan

Dalam tahapan ini, dilakukan perancangan sistem dan pengambilan data menggunakan API Twitter. Pada sistem yang akan dibuat terlebih dahulu flowchart dan DFD untuk mempermudah perancangan sistem.

c. Implementasi

Pada tahapan ini, dibuatlah isi dari menu *user interface* Implementasi Sistem di Shell Python dengan menggunakan bahasa Python, dan melakukan pengambilan data dari Twitter dan melakukan Analisis Sentimen dengan tahapan yaitu *Pre-Processing* meliputi *Case Folding*, *Cleaning Data*, *Tokenization*, *Stemming*, *Normalisasi Bahasa*, dan *Stopwords Removal*. Lalu melakukan metode *TF-IDF* untuk pembobotan kata dalam suatu dokumen.

d. Pengujian

Pengujian di sini meliputi pengujian dari *algoritma* dan *system* yang ada, dimana pengujian *algoritma* yaitu dengan contoh kasus dimuat angka pada sistem tersebut untuk mengetahui apakah system telah berjalan sesuai dengan logikanya atau tidak. Selanjutnya yaitu pengujian system dengan menggunakan *confusion matrix* yang dapat digunakan untuk melakukan akurasi nilai.

1.6 Sistematika Penulisan

Untuk memudahkan dalam mengikuti seluruh uraian dan pembahasan pada penelitian ini, maka penulisan penelitian ini dilakukan dengan sistematika sebagai berikut:

BAB I PENDAHULUAN

Pada bab ini akan menjelaskan mengenai latar belakang penelitian, perumusan penelitian, batasan penelitian, tujuan, manfaat dan sistematika penulisan.

BAB II LANDASAN TEORI

Pada bab ini menjelaskan teori yang digunakan sebagai dasar analisis penelitian. Mencakup tinjauan pustaka meliputi pengertian *Data mining*, *Support Vector Machine (SVM)*, *Analisis Sentimen*, *Pre-processing*, *Term-Frequency-Inverse Document Frequency*, *Confusion Matrix*.

BAB III ANALISI DAN PERANCANGAN

Pada bab ini akan menjelaskan mengenai perancangan dan model cara kerja yang akan dilakukan di dalam studi kasus analisis sentimen terhadap genre musik di twitter.

BAB IV IMPLEMENTASI DAN PEMBAHASAN

Pada bab ini akan menjelaskan *Analisis Sentimen* dari *Pre-Processing Data*, *Tokenization*, *Steaming*, *Normalisasi Bahasa*, *Stopwords Removal*, *TF-IDF*, *Support Vector Machine*, dan *Confusion Matrix*.

BAB V PENUTUP

Bab ini akan menjelaskan kesimpulan yang didapat dari hasil analisis, perancangan dan website serta saran untuk pengembangan lebih baik.