

# BAB 1 PENDAHULUAN

## 1.1. Latar Belakang

Menurut data yang dirilis oleh *Twitter* atau *X* Indonesia pada tahun 2016, 77% dari pengguna media sosial *Twitter* atau *X* di Indonesia adalah pengguna media sosial *Twitter* atau *X* aktif. Hal ini membuat penyebaran informasi melalui media sosial *X* menjadi lebih mudah [1]. Analisis sentimen merupakan bagian dari pemrosesan bahasa alami yang bertugas mengekstrak informasi seperti opini dari dokumen atau teks yang diberikan bisa positif atau negatif. Analisis sentimen ini akan berfokus pada data *X* dapat memberikan wawasan tentang opini masyarakat terhadap berbagai topik, termasuk aplikasi-aplikasi populer seperti aplikasi *Netflix*. Dengan jumlah ulasan atau *tweet* yang sangat banyak pada media sosial *X* bagi pengembang untuk mengetahui sentimen pada ulasan tersebut. Oleh karena itu, diperlukan metode analisis yang efisien dan akurat untuk mengidentifikasi pandangan positif dan negatif dari ulasan pengguna [2].

Berdasarkan latar belakang masalah tersebut maka dilakukan penelitian untuk mengetahui opini masyarakat mengenai aplikasi *Netflix* pada media sosial *X* [1]. Dengan Analisis sentimen pada media sosial *X*, kita dapat mengidentifikasi dan mengklasifikasikan sentimen dari *tweet-tweet* yang ada. Dalam analisis sentiment media sosial *X* ini, menggunakan metode klasifikasi untuk mengklasifikasikan teks menjadi kategori sentimen yang sesuai [3].

Untuk mendapatkan nilai sentimen yang maksimal dengan banyaknya data pada *X* penelitian ini menggunakan ekstraksi fitur dengan *Word2Vec* dan metode klasifikasi yaitu dengan *Support Vector Machine* (SVM) dan *K-Nearest Neighbors* (KNN) [3].

Pada ekstraksi fitur penelitian ini menggunakan *Word2Vec* yang merupakan teknik pemetaan kata ke dalam vektor numerik. Dalam penelitian ini akan menggunakan ekstraksi fitur dengan *Word2Vec* untuk meningkatkan representasi teks sebelum mengaplikasikan metode klasifikasi. *Word2Vec* memungkinkan kita untuk menggabungkan informasi semantik kata-kata dalam representasi vektor [3].

Pada metode klasifikasi menggunakan *Support Vector Machine* (SVM) yang merupakan metode klasifikasi yang membangun model berdasarkan pemisahan optimal antara dua kelas dengan menggunakan *hyperplane* dan mencari *hyperplane* yang memiliki margin maksimum antara dua kelas [3]. Serta menggunakan metode *K-Nearest Neighbors* (KNN) yang merupakan metode klasifikasi berbasis instance yang mengklasifikasikan data berdasarkan mayoritas kelas dari tetangga terdekatnya namun metode ini tidak membangun model, tetapi menghitung jarak antara data uji dengan data latih [2][3].

Menggabungkan klasifikasi SVM atau KNN dengan ekstraksi fitur menggunakan *Word2Vec* memungkinkan peningkatan akurasi dalam analisis sentimen pada data media sosial X mengenai opini masyarakat terhadap aplikasi *Netflix*. Metode ini tidak hanya memfasilitasi identifikasi sentimen positif dan negatif secara lebih tepat, tetapi juga memperdalam pemahaman kita tentang konteks dan nuansa dari tweet yang dianalisis [1]. Oleh karena itu, hasil analisis sentimen ini dapat memberikan wawasan yang lebih kaya bagi pengembang dan pemasar dalam memahami persepsi dan preferensi pengguna terhadap aplikasi *Netflix* [1] [3].

## **1.2. Rumusan Masalah**

Rumusan masalah pada penulisan ini adalah sebagai berikut:

1. Bagaimana cara mengimplementasikan ekstraksi fitur *Word2Vec* untuk merepresentasikan data teks ulasan pengguna aplikasi *Netflix* dalam bentuk vektor yang dapat diproses oleh model *Support Vector Machine* (SVM)?
2. Bagaimana implementasi *Support Vector Machine* (SVM) dapat diterapkan untuk mengklasifikasikan sentimen pada data terkait aplikasi *Netflix*?
3. Bagaimana opini masyarakat terhadap munculnya aplikasi *streaming* film *Netflix*?

## **1.3. Tujuan**

Tujuan penulisan proposal tugas akhir ini adalah sebagai berikut :

1. Mengimplementasikan dan mengkaji pendekatan *Word2Vec* dalam ekstraksi fitur terhadap analisis sentimen.

2. Membandingkan efektivitas dan mengimplementasikan metode *Support Vector Machine* dan *K-Nearest neighbors* (KNN) dalam analisis sentimen.
3. Mengetahui opini masyarakat terhadap aplikasi Netflix.

#### **1.4. Batasan Masalah**

Penelitian ini akan berfokus pada analisis sentimen terhadap data dari media sosial *X* yang berisi opini masyarakat tentang aplikasi *Netflix* tidak mencakup media sosial lainnya. Data yang akan digunakan akan diperoleh dari media sosial *X* dengan menggunakan *Application Program Interface* atau metode *web scraping*. Dataset akan diambil dengan rentang waktu dari Januari 2022 hingga Desember 2024. Penelitian ini akan menggunakan dua metode klasifikasi yaitu *Support Vector Machine* (SVM) dan *K-Nearest Neighbors* (KNN) dan menggunakan *Word2Vec* sebagai ekstraksi fitur dengan data diambil dengan menentukan kata kunci untuk pencarian *tweet* yaitu *Netflix* dengan menggunakan bahasa indonesia. Fokus utama adalah pada analisis sentimen, yaitu mengklasifikasikan teks menjadi kategori sentimen positif dan negatif.

#### **1.5. Rencana Kegiatan**

1. Persiapan dan Rumusan Masalah untuk mempelajari topik yang akan dibahas dan membangun pentingnya topik tersebut sebagai objek penelitian.
2. Studi Literatur
3. Mengumpulkan studi literatur adalah untuk mencari, membaca, dan memahami literatur yang berkaitan.
4. Pengumpulan Data
5. Mengumpulkan dataset melalui *Application Program Interface* dari *X*
6. Melakukan Pra-pemrosesan Data
7. Untuk membersihkan data dari *noise*, menghapus duplikasi, dan mengatasi masalah seperti *emoticon* atau karakter khusus dan melakukan tokenisasi dan normalisasi teks.
8. Pengkodean pada label data

9. Dengan menggunakan klasifikasi SVM dan KNN dan melakukan ekstrasi fitur menggunakan *Word2Vec*.
10. Menganalisis Data
11. Menganalisis data yang sudah didapatkan dan melakukan evaluasi terkait model yang telah di buat.
12. Penulisan Laporan Tugas Akhir
13. Melakukan Penulisan Laporan Tugas Akhir.

## **1.6. Jadwal Kegiatan**

Berdasarkan rencana kegiatan yang telah disusun, maka penulis menyusun jadwal kegiatan sebagaimana dapat dilihat pada Tabel 1.1.

**Tabel 1.1. Tabel Rencana Kegiatan.**

| No. | Kegiatan                       | Bulan |   |   |   |   |   |
|-----|--------------------------------|-------|---|---|---|---|---|
|     |                                | 1     | 2 | 3 | 4 | 5 | 6 |
| 1.  | Persiapan dan Rumusan Masalah  | ■     | ■ | ■ |   |   |   |
| 2.  | Studi Literatur                | ■     | ■ | ■ |   |   |   |
| 3.  | Pengumpulan Data               | ■     | ■ | ■ |   |   |   |
| 4.  | Pra-Pemrosesan Data            |       | ■ | ■ |   |   |   |
| 5.  | Pengkodean Pada Label Data     |       |   | ■ | ■ |   |   |
| 6.  | Analisis Data & Evaluasi Model |       |   |   | ■ | ■ |   |
| 7.  | Penulisan Laporan Tugas Akhir  |       |   |   |   | ■ | ■ |