

# BAB 1

## PENDAHULUAN

### 1.1. Latar Belakang

Ulasan terhadap suatu produk penting bagi para pengembang aplikasi. Ulasan yang terdapat pada *Google Play Store* dapat membantu pengembang menemukan masalah dan area yang perlu ditingkatkan. Selain digunakan untuk mengunduh aplikasi, *Google Play Store* juga menyediakan fitur ulasan aplikasi yang memungkinkan pengguna memberikan komentar. Dengan tersedianya aplikasi Vidio pada *Google Play Store*, pengguna dapat memberikan ulasan atau *review*. Namun, dengan jumlah ulasan yang sangat banyak, sulit bagi pengembang untuk secara manual mengelola dan menganalisis sentimen ulasan tersebut.

Penelitian ini dilakukan untuk menganalisis opini publik mengenai aplikasi Vidio di *Google Play Store*. Mengingat banyaknya data yang terdapat di Internet, data mentah perlu diproses terlebih dahulu agar dapat dimanfaatkan secara optimal. Oleh karena itu, muncullah *opinion mining*, cabang penelitian dari *text mining* [1]. *Opinion mining* adalah teknik *Natural Language Processing* (NLP) yang berfungsi untuk mengidentifikasi nada emosional di balik isi teks. Ini adalah cara populer bagi perusahaan untuk menentukan dan mengategorikan opini tentang suatu produk, layanan, atau ide [2]. Dengan demikian, diperlukan metode analisis yang efisien dan akurat untuk mengidentifikasi pandangan positif atau negatif dari ulasan pengguna.

Studi menggunakan algoritma *Support Vector Machine* (SVM) yang terbukti efektif dalam analisis sentimen. SVM dapat digunakan dalam berbagai kasus seperti untuk *text mining*, mendeteksi wajah, dan bioinformatika [3]. Metode lain seperti *Naïve Bayes* dan *Logistic Regression* juga sering digunakan dalam analisis sentimen. Penelitian sebelumnya telah menggunakan *Naïve Bayes* untuk mengklasifikasikan ulasan pengguna aplikasi Vidio [4]. Hasilnya menunjukkan perbedaan klasifikasi sentimen positif dan negatif berdasarkan kamus pelabelan yang digunakan. Selain itu, metode *K-Nearest Neighbor* (KNN) juga telah dilakukan dalam penelitian ulasan pengguna dengan tingkat akurasi bervariasi [5].

Penerapan metode TF-IDF (*Term Frequency-Inverse Document Frequency*)

digunakan untuk merepresentasikan data ulasan dalam bentuk vektor input. Pendekatan ini membantu memberikan bobot lebih besar pada kata-kata penting dalam teks. Selain itu, data ulasan sering kali mengalami ketidakseimbangan (*imbalance*) antara jumlah ulasan positif dan negatif. Ketidakseimbangan ini dapat memengaruhi performa model klasifikasi. Sehingga, penelitian ini menerapkan teknik SMOTE (*Synthetic Minority Over-sampling Technique*) dalam mengatasi ketidakseimbangan kelas.

Penelitian ini juga akan membandingkan performa beberapa model klasifikasi, termasuk SVM, *Logistic Regression*, dan *Naïve Bayes*, pada data yang telah diseimbangkan menggunakan SMOTE dan data yang tidak diseimbangkan untuk SVM. Pengaruh ratio *oversampling* dalam SMOTE terhadap akurasi klasifikasi juga akan dianalisis. Performa pemodelan akan dievaluasi dengan *confusion matrix*, dan ROC-AUC (*Receiver Operating Characteristic - Area Under Curve*) untuk menganalisis kemampuan model dalam membandingkan kedua kelas.

Penelitian ini diharapkan dapat menawarkan metode yang lebih baik untuk mengelola dan menganalisis sentimen ulasan pengguna aplikasi Vidio di *Google Play Store* secara efisien dan akurat.

## **1.2. Rumusan Masalah**

Rumusan masalah pada penelitian tugas akhir ini adalah sebagai berikut:

1. Bagaimana menilai sentimen publik terhadap aplikasi Vidio pada *Google Play Store* menggunakan *Support Vector Machine*?
2. Bagaimana performansi *Support Vector Machine* dalam mengidentifikasi sentimen ulasan aplikasi Vidio?

## **1.3. Tujuan**

Tujuan penulisan penelitian tugas akhir ini adalah sebagai berikut:

1. Mengimplementasikan *Support Vector Machine* dalam analisis sentimen ulasan aplikasi Vidio.
2. Menguji tingkat performansi *Support Vector Machine* dalam mengidentifikasi sentimen ulasan aplikasi Vidio.

#### 1.4. Batasan Masalah

1. Dataset yang digunakan bersumber dari data *review* ulasan pengguna aplikasi Vidio di *Google Play Store* dengan kurun waktu dari bulan September 2018 hingga Januari 2023.
2. Dataset yang digunakan berbahasa Indonesia dengan format .CSV.

#### 1.5. Jadwal Pelaksanaan

Berdasarkan rencana kegiatan yang telah disusun, maka penulis menyusun jadwal kegiatan sebagaimana dapat dilihat pada Tabel 1.

Tabel 1. Jadwal Pelaksanaan Tugas Akhir

No.	Deskripsi Tahapan	Bulan 1	Bulan 2	Bulan 3	Bulan 4	Bulan 5	Bulan 6
1	Persiapan dan Perumusan Masalah						
2	Studi Literatur dan Kajian Pustaka						
3	Pengumpulan Data						
4	Analisis dan Perancangan Sistem						
5	Implementasi Sistem						
6	Analisis Hasil Implementasi						
7	Penyusunan Laporan/Buku TA						