

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

Perkembangan teknologi digital telah mengubah perilaku masyarakat dalam berbelanja, terutama dengan meningkatnya penggunaan platform *e-commerce* seperti Shopee. Pengguna aplikasi ini sering memberikan ulasan mengenai pengalaman berbelanja mereka, yang berisi opini dan emosi terhadap layanan atau produk. Informasi tersebut penting bagi perusahaan untuk menilai kualitas layanan dan kepuasan pengguna. Namun, jumlah ulasan yang besar membuat analisis manual menjadi tidak efisien.

Pendekatan *Natural Language Processing (NLP)* digunakan untuk mengotomatisasi pemahaman teks dan opini pengguna [1]. Salah satu penerapan NLP adalah *analisis sentimen*, yang bertujuan mengidentifikasi opini positif, negatif, atau netral dari teks [2]. Berbagai penelitian terdahulu telah menerapkan metode *Naïve Bayes* dan menunjukkan hasil akurasi yang baik untuk analisis sentimen [3]. Penelitian lain juga menggunakan *Naïve Bayes* dalam klasifikasi opini publik di media sosial dengan hasil yang signifikan [4]. Selain itu, *Support Vector Machine (SVM)* banyak digunakan karena mampu menangani data berdimensi tinggi dengan hasil klasifikasi yang stabil [5]. Penerapan *SVM* dengan pembobotan TF-IDF terbukti memberikan akurasi hingga 89,05% pada analisis sentimen aplikasi *PeduliLindungi* [6]. Metode ini juga terbukti unggul dalam mengklasifikasikan ulasan *e-commerce* dan aplikasi digital [7].

Meski demikian, metode-metode tersebut masih memiliki keterbatasan dalam memahami konteks semantik dan menangani kata baru (*out-of-vocabulary*) [8]. Bahasa Indonesia yang memiliki banyak bentuk tidak baku dan gaya informal menambah kompleksitas pemrosesan teks [9]. Untuk mengatasi hal ini, digunakan metode *FastText* yang dikembangkan oleh *Facebook AI Research* [10]. *FastText*

mempelajari representasi kata berdasarkan unit sub-kata (*n-gram*) sehingga mampu mengenali kata baru dan kesalahan ejaan [11]. Keunggulan tersebut membuat *FastText* efektif dalam pengolahan bahasa alami, termasuk untuk teks berbahasa Indonesia [12].

Pada penelitian [13] telah dibuat analisis sentimen dengan menggunakan metode *Support Vector Machine* (SVM). Dataset yang digunakan didapatkan dengan menggunakan *web scraping*. Hasil penelitian pada Bag of Words menggunakan metode SVM mendapatkan accuracy sebesar 47%. Hasil penelitian pada TF-idf menggunakan metode SVM mendapatkan accuracy sebesar 71%. Hasil penelitian pada *FastText* menggunakan metode SVM mendapatkan accuracy sebesar 49%.

Pada penelitian [14] telah dibuat analisis sentimen dengan menggunakan metode *Support Vector Machine* (SVM). Dataset yang digunakan didapatkan dengan menggunakan *web scraping*. Hasil penelitian pada *FastText* menggunakan metode SVM mendapatkan accuracy sebesar 88%. Hasil penelitian pada *Word2vec* menggunakan metode SVM mendapatkan accuracy sebesar 76%.

Pada penelitian [15] telah dibuat analisis sentimen dengan menggunakan metode *Support Vector Machine* (SVM). Dataset yang digunakan didapatkan dengan menggunakan *web scraping*. Hasil penelitian pada menggunakan metode SVM pada website Ruang Guru mendapatkan accuracy sebesar 99%. Hasil penelitian pada menggunakan metode SVM pada website Zenius mendapatkan accuracy sebesar 96%. Hasil penelitian pada menggunakan metode SVM pada website Quipper mendapatkan accuracy sebesar 82%.

Berdasarkan uraian tersebut, penelitian ini berfokus pada penerapan *FastText* sebagai metode *word embedding* dan *Support Vector Machine* (SVM) sebagai algoritma klasifikasi utama dalam analisis sentimen ulasan pengguna aplikasi Shopee.

## 1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, terdapat beberapa permasalahan utama yang perlu diteliti dalam analisis sentimen ulasan aplikasi Shopee di Google Play Store menggunakan metode *FastText*. Adapun rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana proses analisis sentimen ulasan aplikasi menggunakan representasi fitur berbasis *FastText*?
2. Bagaimana performa metode Support Vector Machine (SVM) dalam melakukan klasifikasi sentimen dengan fitur yang dihasilkan oleh *FastText*?
3. Kernel SVM manakah yang menghasilkan performa terbaik dalam klasifikasi sentimen ulasan aplikasi berbasis fitur *FastText*?

## 1.3. Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Menerapkan *FastText* untuk membentuk representasi fitur teks pada ulasan aplikasi.
2. Menganalisis performa Support Vector Machine (SVM) dalam mengklasifikasikan sentimen ulasan aplikasi menggunakan fitur *FastText*.
3. Mengetahui kernel SVM yang menghasilkan performa terbaik dalam klasifikasi sentimen ulasan aplikasi.

## 1.4. Batasan Masalah

Pada penelitian ini terdapat batasan-batasan masalah agar penelitian yang dilakukan tidak menyimpang dari pokok pembahasan yang diangkat, berikut batasan masalah dari penelitian ini:

1. Data yang digunakan dalam penelitian ini berupa ulasan teks pengguna aplikasi yang telah memiliki label sentimen.
2. Proses praproses data teks dibatasi pada case folding, pembersihan karakter, tokenisasi, dan penghapusan stopword.
3. FastText digunakan untuk membentuk representasi fitur teks, bukan sebagai metode klasifikasi mandiri.
4. Support Vector Machine (SVM) digunakan sebagai metode klasifikasi utama dalam analisis sentimen, dengan pengujian menggunakan kernel Linear dan RBF.
5. Evaluasi performa model klasifikasi dilakukan menggunakan confusion matrix, serta metrik accuracy, precision, recall, dan F1-score.
6. Penelitian ini tidak membahas aspek optimasi hiperparameter secara mendalam maupun penerapan metode pembelajaran mendalam (deep learning) lainnya.

