

## BAB I PENDAHULUAN

### 1.1 Latar Belakang

DANA merupakan aplikasi dompet digital yang memiliki berbagai fitur sehingga mempermudah banyak sekali transaksi bagi penggunanya. DANA dapat di unduh pada *Google Play store* bagi pengguna android dan *App store* bagi pengguna iphone. *Google Play* memberikan fitur ulasan untuk setiap aplikasi yang tersedia di dalamnya sehingga para pengguna dapat memberikan komentar atau ulasan mengenai kepuasan maupun kritik dan saran mereka terhadap aplikasi yang mereka unduh. Tujuannya adalah untuk mengevaluasi kinerja aplikasi dari developer sehingga dapat dilakukan perbaikan dan penyesuaian aplikasi agar dapat sesuai dengan keinginan pengguna. Komentar yang dibuat pun beragam sehingga bisa digunakan sebagai data sampel. Kebanyakan pengguna hanya melihat *rating* untuk dijadikan aplikasi tersebut bagus atau tidak, namun sebenarnya *rating* tidak dapat dijadikan sebagai acuan karena ada beberapa ulasan yang sifatnya negatif namun *rating*nya tinggi sehingga tidak sesuai dengan ulasan yang tertulis. Menyebabkan beberapa kebingungan atau keambiguan tentang hal mana yang benar adanya. Oleh sebab itu diperlukan suatu teknik untuk mengorganisir dan mengelola data teks untuk mengetahui kebenaran data dengan menggunakan analisis sentimen.

Analisis sentimen adalah studi komputasi tentang opini, perilaku dan emosi entitas. Entitas-entitas inilah yang dapat menggambarkan perilaku seseorang dalam membentuk suatu informasi tentang peristiwa. Analisis sentimen merupakan proses mengekstraksi, memahami dan memproses data teks untuk mendapatkan informasi emosional yang terkandung dalam kalimat opini. Analisis sentimen juga dapat dilakukan untuk mengetahui opini atau komentar terhadap suatu entitas, apakah memiliki sentimen positif atau negatif sehingga dapat dijadikan acuan untuk memperbaiki layanan dan kualitas sebuah produk (Shiddieqy, dkk., 2016).

Pada penelitian sebelumnya yang dilakukan oleh Romadoni (2020), dengan judul *Text Mining* untuk analisis sentimen pelanggan terhadap layanan uang elektronik menggunakan algoritma *Support Vector Machine (SVM)* ada banyak

metode yang dapat digunakan untuk mengklasifikasi seperti algoritma *Support vector machine*. *Support Vector Machine (SVM)* merupakan metode klasifikasi yang dapat digunakan dalam permasalahan linier dan non-linier. Metode ini memiliki kelebihan dalam mempresentasikan contoh kasus seperti titik dalam ruang, dipetakan sehingga contoh dari kategori dipisahkan oleh margin yang jelas seoptimal mungkin. Hal ini memberikan hasil terbaik daripada algoritma *Naive Bayes* dan *tool* klasifikasi sentimen lainnya. Pada penelitian ini memperoleh nilai akurasi sebesar 98,7% dengan 3467 data set (Romadoni, dkk., 2020). Penelitian selanjutnya dilakukan oleh Nuraeni Herlinawati yang mengklasifikasi ulasan aplikasi *Zoom Cloud Meetings* menggunakan metode *Naive Bayes* dan *Support Vector Machine (SVM)* memperoleh nilai akurasi untuk *Naive Bayes* sebesar 74,37% sedangkan nilai akurasi untuk metode *Support Vector Machine (SVM)* sebesar 81,22%. Dalam penelitian tersebut dapat kita lihat dan ketahui bahwa nilai akurasi yang didapatkan oleh metode *Support Vector Machine (SVM)* lebih besar 6,85% dibandingkan dengan metode *Naive Bayes* dengan 1.007 record dataset (Herlinawati, dkk., 2020). Dengan penjelasan penelitian terdahulu tersebut maka metode *Support Vector Machine (SVM)* ini sangat cocok untuk digunakan karena metode *Support Vector Machine (SVM)* memiliki tingkat akurasi yang sangat tinggi dibandingkan dengan metode klasifikasi *Naive Bayes* dan dapat digunakan dalam permasalahan yang sederhana maupun dengan dimensinya yang lebih tinggi.

Pada ulasan aplikasi di *Google Play Store* memberikan kebebasan memberi komentar, masukan maupun kritikan oleh para pengguna. Penggunaan bahasa yang bebas pun tak dapat terelakkan dalam pengulasan sebuah aplikasi. Banyak dari para pengguna yang menyelipkan kata-kata tidak baku, penyingkatan kata, dan kesalahan ejaan atau *typo*. Pada tahapan *preprocessing text* juga tidak dapat mengatasi semua variasi unik penggunaan bahasa seperti kesalahan eja dan penyingkatan kata tersebut. Maka dari itu diperlukan suatu algoritma yang dapat memberikan perbaikan kata maupun kesalahan ejaan pada ulasan para pengguna, metode yang dapat digunakan yaitu *Levenshtein distance*. Metode *Levenshtein Distance* merupakan salah satu metode yang dapat digunakan sebagai solusi

dalam mengatasi kesalahan ejaan dengan menghitung jumlah perbedaan *string*, antara *string* target dan *string input* menggunakan matriks. Pada penelitian sebelumnya yang dilakukan oleh Ferly Gunawan dkk (2017), yang berjudul “Analisis Sentimen Pada Ulasan Aplikasi *Mobile* Menggunakan *Naive Bayes* dan Normalisasi Kata Berbasis *Levenshtein Distance* (Studi Kasus Aplikasi *BCA Mobile*)” dijelaskan bahwa *Levenshtein Distance* dapat diterapkan untuk normalisasi kata pada ulasan aplikasi *mobile* dengan studi kasus ulasan aplikasi *BCA mobile* (Gunawan, Fauzi, dan Adikara 2017).

Dari penjelasan tersebut maka dapat dilakukan sebuah kolaborasi antara metode klasifikasi *Support Vector Machine (SVM)* dengan algoritma *Levenshtein distance*. Kolaborasi ini bertujuan untuk memperbaiki beberapa kesalahan ejaan dan perbaikan kata baku pada setiap ulasan aplikasi *DANA* yang akan digunakan dan nantinya diklasifikasi oleh metode *Support vector machine*. Pada penggabungan dua metode ini diharap memberikan hasil akurasi yang lebih baik dari penelitian terdahulu.

Berdasarkan semua penjelasan yang penulis paparkan maka penulis akan melakukan penelitian pada data ulasan oleh pengguna aplikasi *DANA* yang diklasifikasi menggunakan metode *Support Vector achine* dengan perbaikan kata tidak baku dan perbaikan kesalahan ejaan menggunakan *Levenshtein distance*. Berdasarkan latar belakang tersebut, maka penulis akan melakukan penelitian dengan judul “**Penerapan *Levenshtein Distance* Terhadap *Support Vector Machine (SVM)* Dalam Analisis Sentimen Teks (Studi Kasus : Ulasan Aplikasi *Dana*)**”.

## 1.2 Rumusan Masalah

Berdasarkan uraian permasalahan yang ada pada latar belakang di atas, maka rumusan masalah pada skripsi ini adalah:

Berapa hasil perbdaningan nilai akurasi yang diperoleh jika menggunakan normalisasi kata *Levenshtein Distance* terhadap metode *Support Vector Machine (SVM)* dan tanpa menggunakan metode *Levenshtein Distance* terhadap Analisis sentimen pada ulasan pengguna aplikasi *DANA*?

### 1.3 Tujuan Penelitian

Tujuan penelitian berdasarkan permasalahan yang telah diidentifikasi di atas adalah:

Mengetahui kinerja *Levenshtein Distance* terhadap metode *Support Vector Machine (SVM)* dengan menggunakan nilai tingkat akurasi yang didapatkan pada klasifikasi sentimen data ulasan aplikasi DANA di *Google Play Store*.

### 1.4 Manfaat Penelitian

Penelitian ini diharapkan memiliki manfaat baik secara teoritis maupun praktis di bidang penelitian dan aplikasinya pada masyarakat luas. Berdasarkan tujuan di atas, adapun manfaat dari penelitian ini adalah sebagai berikut:

1. Diharapkan mampu menjadi referensi atau sumbangan pemikiran bagi peneliti lain dalam bidang analisis sentimen secara umum dengan menggunakan algoritma *Support Vector Machine (SVM)* dan *Levenshtein Distance*.
2. Dapat menambah pengetahuan peneliti khususnya dibidang pengklasifikasian algoritma *Support Vector Machine (SVM)* dan *Levenshtein Distance*.

### 1.5 Batasan Penelitian

Penelitian ini memiliki ruang lingkup dan batasan-batasan dari pembahasan permasalahan serta topik yang berfungsi untuk memfokuskan penelitian, agar permasalahan pada penelitian ini tidak menyimpang dan lebih terarah pada titik penelitian yang dilakukan. Berikut adalah batasan-batasan masalah yang dilakukan di penelitian ini :

1. Data sampel yang digunakan pada penelitian ini diunduh dari ulasan para pengguna aplikasi DANA di *Google Play store* pada bulan Januari hingga Maret tahun 2023 sebanyak 1000 data ulasan.
2. Hasil klasifikasi pada penelitian ini dibagi menjadi 2 kelas yaitu sentimen positif dan sentimen negatif.