

BAB I

PENDAHULUAN

1.1 Latar Belakang

Analisis sentimen merupakan salah satu teknik dari cabang penelitian *Text Mining* yang digunakan dalam mengklasifikasi dokumen atau data teks berupa opini berdasarkan sentimen [1]. Analisis sentimen telah banyak digunakan dalam *review* produk yang biasanya berisi pendapat konsumen terhadap suatu produk. Tidak hanya itu, teknik ini juga dapat digunakan untuk mengkategorikan opini positif dan negatif terhadap suatu ulasan, baik dari program pemerintah maupun suatu aplikasi yang dirasa sering digunakan masyarakat pada saat ini. Analisis yang dilakukan bertujuan untuk mengklasifikasi ulasan para pengguna terhadap kenyamanan maupun keluhan sistem dari aplikasi melalui teknik *text mining* dengan data yang diperoleh dari berbagai sumber, misal dari *twitter*, *instagram*, maupun *website google play store* yang biasa digunakan untuk mengunduh suatu aplikasi yang terdapat pada perangkat *android* [2]. Mengetahui faktor keluhan yang paling banyak dirasakan oleh pengguna terhadap suatu aplikasi dapat dijadikan sebagai masukan untuk perbaikan sistem kedepannya.

Pada zaman serba digital ini, kehadiran aplikasi yang memudahkan masyarakat sangat dibutuhkan, salah satunya yaitu dompet digital atau mungkin lebih dikenal dengan *E-wallet*. Dompet digital merupakan layanan keuangan berbentuk aplikasi yang digunakan untuk melakukan transaksi dalam berbagai keperluan dengan saldo. Dompet digital menjadi salah satu solusi pembayaran yang sering digunakan oleh masyarakat terutama para pemuda. Aplikasi tersebut memberikan banyak kemudahan dalam layanan transaksi *online*, diantaranya dapat menyimpan uang dalam bentuk elektronik, mengirim, dan juga membayar berbagai keperluan transaksi[2].

Menurut *Kompas.com*, aplikasi dompet digital atau *e-wallet* yang umumnya digunakan oleh masyarakat Indonesia ada 10. Namun, disini peneliti mengambil 5 aplikasi dengan peringkat teratas yang sering digunakan dengan survei dari sumber terpercaya. Kelima dompet digital dengan peringkat teratas

pada tahun 2022 menurut *Kompas.com*, yaitu Gopay (88%), Dana (83%), OVO (79%), ShopeePay (76%), dan LinkAja (30%) [3]. Tentunya kelima dompet digital tersebut memiliki keunggulan masing-masing sehingga dipercaya oleh sebagian besar masyarakat.

Namun, disamping tingginya minat masyarakat terhadap aplikasi tersebut, pastinya *developer* harus selalu meningkatkan layanan agar bisa memuaskan pengguna tanpa adanya kendala. Kelima dompet digital diatas, sudah cukup mewakili dan dimaksudkan agar nantinya bisa menjadi bahan evaluasi aplikasi lainnya. Seperti yang dijelaskan pada paragraf sebelumnya, analisis sentimen menjadi solusi dalam mengklasifikasikan ulasan dari pengguna terhadap layanan aplikasi. Nantinya akan diperoleh sentimen dengan opini positif ataupun negatif dari pengguna yang nantinya bermanfaat untuk perbaikan aplikasi.

Menurut *kompasiana.com*, *Google Play Store* adalah *marketplace digital* perusahaan yang secara *default* sudah terinstall di *android* dan menjadi tempat pengguna untuk mengunduh berbagai aplikasi [4]. Pada *Google Play Store*, pengguna akan sering membaca *review* pengguna lain serta reputasi aplikasi sebelum mengunduh [5]. Hal itu membuat analisis ulasan pengguna sangat menarik dan bermanfaat untuk *developer* dalam meningkatkan layanan aplikasi [5]. Maka dari itu, proses *crawling* akan dilakukan pada website *google play store* dengan *crawling* data ulasan pengguna pada lima aplikasi yang telah dijelaskan di paragraf sebelumnya. Akan dilakukan *crawling* sebanyak 400 ulasan pada masing-masing aplikasi tersebut, yang nantinya diperoleh total 2000 ulasan. Klasifikasi sentimen ulasan ini menggunakan metode *Support Vector Machine* (SVM).

SVM adalah salah satu metode klasifikasi *supervised learning* yang memprediksi kelas berdasarkan pola dari hasil proses pelatihan yang dicetuskan oleh Vladimir Vapnik tahun 1992. Klasifikasi dilakukan dengan *hyperplane* yang memisahkan antara kelas positif dan negatif [2]. *Hyperplane* menjadi pemisah terbaik antara kedua kelas yang dapat dilakukan dengan mengukur margin *hyperplane* dan mencari titik maksimalnya.

Performa klasifikasi tidak maksimal apabila ditemukan *imbalance class* pada data. *Imbalance class* sering ditemukan pada topik analisis sentimen dan nantinya akan mengakibatkan hasil prediksi dari model *machine learning* lebih cenderung pada data dengan kelas yang dominan [6]. Oleh karena itu dalam mengatasi *imbalance* tersebut, terdapat beberapa metode yang membantu mengurangi kelemahan dari proses diatas, salah satunya adalah SMOTE. SMOTE adalah metode *oversampling* yang digunakan untuk mengatasi masalah *imbalance class* pada data. Menurut penelitian, umumnya metode ini lebih akurat daripada pendekatan *oversampling* biasa lainnya [7].

Pada penelitian terdahulu, dilakukan analisis sentimen tentang pendapat pro dan kontra masyarakat mengenai vaksinasi *Booster Covid-19* pada tahun 2022 yang dilakukan pengambilan data dari *twitter*. Dilakukan perbandingan terhadap tiga metode klasifikasi, dan yang menghasilkan akurasi paling tinggi adalah metode SVM dengan nilai 83.33%, *Naïve Bayes* sebesar 70% dan *Decision Tree* 79% [6]. Pada tahun 2022 pula dilakukan penelitian yang melakukan perbandingan metode SVM dan *Naïve Bayes* dengan SMOTE dalam analisa sentimen vaksinasi *covid-19*. Dalam penelitian tersebut disimpulkan bahwa hasil pengujian menggunakan SVM dengan SMOTE memperoleh nilai akurasi lebih baik yaitu 70.51%, sedangkan metode SVM tanpa menerapkan SMOTE diperoleh nilai 63.13% [7].

Berdasarkan penelitian diatas, menunjukkan bahwa dengan menggunakan SVM pada klasifikasi akan menghasilkan nilai akurasi yang lebih baik daripada metode *naïve bayes* dan *decision tree*. Penambahan metode dengan SMOTE juga diharapkan dapat membantu meningkatkan hasil yang optimal terhadap klasifikasi dan nilai akurasi. Maka dari itu, penelitian ini dilakukan penerapan SVM dan SMOTE pada analisis sentimen dompet digital berdasarkan ulasan pengguna.

1.2 Perumusan Masalah

1.2.1 Permasalahan

Permasalahan yang terdapat pada penelitian ini yaitu:

Kurangnya penelitian yang melakukan klasifikasi sentimen mengenai ulasan dompet digital pada *google play store* menggunakan metode SVM untuk mengetahui nilai akurasi yang diperoleh, serta terjadinya *imbalance data* pada dataset yang menyebabkan kesalahan klasifikasi kelas minoritas dan mayoritas sehingga memberikan performa klasifikasi yang kurang maksimal pada akurasi.

1.2.2 Metode Usulan

Berdasarkan permasalahan, metode yang diusulkan pada penelitian ini adalah *Support Vector Machine* (SVM) sebagai algoritma klasifikasi teks pada ulasan dan metode SMOTE untuk mengatasi *imbalance data*.

1.2.3 Pertanyaan Penelitian

Berdasarkan uraian permasalahan dan solusi yang akan digunakan untuk klasifikasi sentimen didapatkan pertanyaan sebagai berikut:

1. Berapa nilai akurasi yang didapatkan untuk klasifikasi sentimen pada ulasan dompet digital dengan menggunakan metode SVM?
2. Bagaimana pengaruh SMOTE pada metode SVM terhadap nilai akurasi sistem dalam penelitian?

1.3 Tujuan dan Manfaat

1.3.1 Tujuan Penelitian

Berdasarkan latar belakang sebelumnya, tujuan dari penelitian ini adalah sebagai berikut:

1. Penerapan metode SVM bertujuan untuk memperoleh hasil akurasi yang baik dalam melakukan klasifikasi ulasan positif dan negatif.
2. Menerapkan penggunaan SMOTE sebagai metode *imbalance data* dalam klasifikasi sentimen menggunakan SVM dan membandingkan hasil akurasi yang didapat.

1.3.2 Manfaat Penelitian

Manfaat yang diharapkan dari klasifikasi sentimen pada penelitian ini adalah sebagai berikut:

1. Mengetahui performa metode SVM dalam melakukan klasifikasi teks.
2. Mengetahui seberapa besar pengaruh SMOTE pada metode SVM terhadap nilai akurasi.
3. Sebagai literasi untuk penelitian yang berhubungan dengan analisis sentimen pada dompet digital, metode SVM, dan penerapan SMOTE.

1.4 Batasan Masalah

Penelitian ini hanya dilakukan untuk klasifikasi sentimen pada dompet digital yang dibatasi dengan beberapa hal sebagai berikut:

1. Ulasan diambil dari web *google play store* dengan ulasan bahasa Indonesia.
2. Ulasan mengambil dari lima aplikasi dompet digital teratas yang sering digunakan masyarakat, yaitu Gopay, Dana, OVO, ShopeePay, dan LinkAja. (diambil dari sumber berita *Kompas.com*)
3. Data yang diambil 400 ulasan pada masing-masing aplikasi.
4. Data yang diambil diperoleh pada bulan September - November 2022, dengan total 2000 ulasan dari lima aplikasi.

1.5 Sistematika Penulisan

Sistematika penulisan penelitian ini adalah sebagai berikut:

1. Bab I Pendahuluan
Bab ini berisi penjelasan tentang latar belakang penelitian, permasalahan, tujuan dan manfaat penelitian, batasan masalah, dan sistematika penulisan.
2. Bab II Kajian Pustaka
Bab ini berisi teori-teori sebagai bahan penunjang yang berkaitan dengan penelitian.
3. Bab III Metode Usulan

Bab ini menjelaskan langkah-langkah yang digunakan untuk menyelesaikan permasalahan menggunakan metode yang ditentukan.

4. Bab IV Hasil dan Pembahasan

Bab ini menjelaskan lebih rinci mengenai penerapan dan hasil klasifikasi sentimen yang didapatkan menggunakan metode SVM (*Support Vector Machine*) serta penerapan penggunaan SMOTE dan pengaruhnya pada hasil yang didapatkan.

5. Bab V Penutup

Bab ini berisi kesimpulan yang didapatkan dari penelitian yang telah dilakukan serta saran untuk pengembangan penelitian selanjutnya.