

BAB I

PENDAHULUAN

1.1 Latar Belakang

Setiap aplikasi umumnya selalu memiliki kelebihan dan kekurangan yang dapat memunculkan beberapa reaksi pengguna terhadap aplikasi tersebut. Kolom ulasan menjadi salah satu tempat di Google Play Store bagi pengguna untuk dapat memberikan reaksi atau *feedback* terhadap suatu aplikasi (Giovani et al., 2020). Terdapat beberapa aplikasi dengan jumlah unduhan dan pemberian *rating* cukup tinggi, salah satunya Google Classroom. Google Classroom adalah aplikasi manajemen sistem pembelajaran yang dikembangkan oleh Google untuk digunakan sebagai evaluasi tugas yang diserahkan (Gaib, Nanja and Dalai, 2021).

Google Classroom dipilih karena merupakan aplikasi manajemen sistem pembelajaran yang paling banyak digunakan menurut Lembaga Arus Survei Indonesia yang dikutip oleh Kompas.com. Survei yang dilakukan pada Oktober 2020 di 34 provinsi di Indonesia, Google Classroom terpilih sebagai platform yang paling banyak digunakan untuk melakukan pembelajaran jarak jauh (PJJ), dengan 26,1% pengguna lebih memilih Ruang Guru (17,1%) dan Ruang Belajar (15,2%) (Kamil, 2020). Ulasan pada aplikasi Google Classroom harus diperhatikan karena banyak aplikasi pembelajaran *online* lainnya dengan jumlah unduhan dan *rating* yang hampir sama dengan Google Classroom, sehingga predikat terbaik untuk sebuah aplikasi menjadi tidak relevan (Erfina et al., 2020). Selain itu, mayoritas ulasan yang diberikan justru tidak sesuai dengan *rating* yang diberikan, sehingga

ulasan dan *rating* tersebut tidak dapat dijadikan acuan sebagai penilaian dari suatu aplikasi.

Terdapat suatu analisis data melalui proses komputasi yang bertujuan untuk memperoleh informasi tentang sentimen dari suatu aplikasi yang disebut dengan analisis sentimen. Tujuan dilakukan analisis sentimen yaitu untuk memberikan informasi yang bernilai kepada seseorang yang terkandung dalam kumpulan data yang tidak terstruktur (Herlinawati et al., 2020). Proses analisis sentimen memiliki beberapa tahapan, salah satunya pelabelan. Pelabelan adalah tahap untuk memberikan label apakah suatu ulasan memiliki nilai positif, negatif atau netral (Ilmawan and Mude, 2020). Salah satu teknik pelabelan yang dapat digunakan untuk memberikan label terhadap data ulasan yaitu pelabelan berdasarkan kamus leksikon. Metode pelabelan *Valence Aware Dictionary and sEntiment Reasoner* (VADER) merupakan salah satu metode yang termasuk pada metode pelabelan berdasarkan kamus leksikon. Kelebihan dari metode pelabelan VADER yaitu memiliki kamus yang berisi nilai polaritas dari setiap kata (Abimanyu et al., 2022).

Penelitian yang dilakukan oleh (Susilo, 2020) menuturkan proses pelabelan data dilakukan dengan membandingkan kedua metode pelabelan, metode pertama dilakukan dengan mengubah *rating* yang telah diberikan oleh pengguna menjadi sentimen positif, netral dan negatif. Metode kedua dilakukan pelabelan menggunakan kamus *lexicon* untuk melakukan pelabelan. Pada kamus *lexicon* terdapat bobot kata yang menunjukkan kata tersebut termasuk dalam kategori positif atau negatif, dengan melihat nilai pada bobot tersebut.

Penelitian terkait sebelumnya menghasilkan metode TF-IDF dan SVM dapat diterapkan pada proses klasifikasi dengan hasil pengukuran yang cukup baik dengan tingkat akurasi 83,3%, namun hasil yang diperoleh tidak lebih unggul dari penelitian sebelumnya. Hal ini disebabkan perbedaan *dataset*, proses *labeling*, tahapan *preprocessing* dan penggunaan fitur. Selain itu, pengaruh TF-IDF sebagai ekstraksi fitur terhadap pengukuran kinerja model tidak begitu besar, namun sebaiknya menggunakan pembobotan TF-IDF (Fransiska, Rianto and Irham Gufroni, 2020). Selain itu, penelitian terkait lainnya yang menghasilkan performa klasifikasi metode VADER yang memiliki akurasi sebesar 77%, sedangkan *TextBlob* memiliki akurasi 74% dan *SentiWordNet* memiliki akurasi 62% (Bonta, Kumaresh and Janardhan, 2019).

Berdasarkan latar belakang permasalahan, penelitian ini berfokus tentang bagaimana penerapan dan pengaruh pelabelan VADER pada algoritma klasifikasi *Support Vector Machine* terhadap proses analisis sentimen. Oleh karena itu, penelitian ini mengambil judul “Optimalisasi *Support Vector Machine* (SVM) Menggunakan Pelabelan VADER Pada Analisis Sentimen Ulasan Google Classroom”.

1.2 Rumusan Masalah

Berdasarkan latar belakang permasalahan, maka dapat dirumuskan masalah dalam penelitian ini yaitu bagaimana performa algoritma *Support Vector Machine* (SVM) dalam melakukan analisis sentimen ulasan Google Classroom dengan menerapkan metode pelabelan VADER.

1.3 Tujuan Penelitian

Tujuan penelitian ini adalah untuk menganalisis performa algoritma *Support Vector Machine* (SVM) dalam melakukan analisis sentimen pada ulasan Google Classroom dengan menerapkan metode pelabelan VADER.

1.4 Manfaat Penelitian

Penelitian ini diharapkan dapat bermanfaat bagi seluruh pihak yang terkait, diantaranya:

1. Dapat mengetahui performa algoritma *Support Vector Machine* (SVM) dalam melakukan analisis sentimen ulasan Google Classroom dengan menerapkan metode pelabelan VADER.
2. Dapat mengetahui pengaruh metode pelabelan VADER pada performa algoritma *Support Vector Machine* (SVM).

1.5 Batasan Masalah

Terdapat beberapa batasan masalah yang digunakan pada penelitian ini, sebagai berikut:

1. Sumber data berasal dari data primer yaitu *dataset* yang bersumber dari ulasan Google Play Store yang dihasilkan melalui metode *web scraping* menggunakan *library* dari *python* yaitu *google-play-scraper*.
2. Data yang diambil berjumlah 10.000 ulasan tekstual berbahasa inggris.
3. Proses pelabelan data ulasan menggunakan kamus leksikon dari VADER.
4. Proses klasifikasi data menggunakan algoritma *Support Vector Machine* (SVM).

5. Proses *data preprocessing* yang dilakukan tidak menggunakan penanganan negasi karena metode pelabelan VADER dapat menangani secara langsung terhadap kata yang mengandung negasi.

1.6 Metodologi Penelitian

Metode penelitian yang digunakan pada penelitian ini yaitu metode kualitatif. Adapun tahapan penelitian yang dilakukan untuk menyelesaikan penelitian ini adalah sebagai berikut:

1. Studi Literatur

Tahap ini dilakukan dengan eksplorasi terhadap beberapa penelitian terkait berdasarkan topik penelitian untuk mengidentifikasi permasalahan penelitian.

2. Pengumpulan Data

Tahap ini dilakukan dengan mengumpulkan data yang digunakan untuk penelitian. Metode pengumpulan data yang dilakukan melalui metode *web scraping*.

3. *Data Preprocessing*

Tahap ini dilakukan untuk membersihkan data dengan beberapa proses diantaranya *Case Folding*, *Tokenizing*, *Filtering/Stopword Removal* dan *Lemmatization*.

4. Pelabelan Data

Tahap ini dilakukan untuk memberikan label dengan kelas sentimen dengan metode pelabelan VADER.

5. Pembobotan Kata

Tahap ini dilakukan untuk menentukan bobot kata pada setiap data ulasan dengan metode TF-IDF.

6. Klasifikasi Data

Tahapan ini dilakukan dengan mengklasifikasikan data dengan memprediksi sentimen pada setiap data ulasan menggunakan algoritma SVM.

7. Evaluasi Hasil

Tahap ini dilakukan dengan mengevaluasi performa atau tingkat akurasi algoritma klasifikasi menggunakan *confusion matrix* dan hasil klasifikasi dan pelabelan dengan visualisasi.

1.7 Sistematika Pembahasan

Pembahasan pada penelitian ini terbagi menjadi lima BAB, masing-masing BAB diuraikan sebagai berikut:

BAB I PENDAHULUAN

Bab ini membahas mengenai latar belakang, rumusan masalah, tujuan dan manfaat penelitian, metodologi penelitian, serta sistematika pembahasan pada penelitian yang dilakukan.

BAB II TINJAUAN PUSTAKA

Bab ini menguraikan teori-teori yang berhubungan dengan penelitian analisis sentimen yang menjadi pokok utama penelitian serta *state of the art* dari penelitian sebelumnya.

BAB III METODOLOGI PENELITIAN

Bab ini berisi deskripsi secara umum terhadap penelitian yang dilakukan dan penjelasan tentang apa saja yang dilakukan pada penelitian ini.

BAB IV HASIL DAN PEMBAHASAN

Bab ini berisi tentang implementasi dari bab sebelumnya, khususnya implementasi sesuai dengan metodologi penelitian yang digunakan serta pengujian untuk mendapatkan hasil yang sesuai dengan tujuan penelitian.

BAB V KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan dan saran berdasarkan temuan penelitian serta garis besar metode penelitian yang digunakan. Kesimpulan merupakan hasil akhir dari penelitian, sedangkan saran meliputi rekomendasi yang sesuai dengan keterbatasan sistem yang direkomendasikan.