

BAB V

PENUTUP

Bab ini berisi kesimpulan dari penelitian yang telah dilakukan sesuai dengan tahapan metodologi yang telah dijelaskan sebelumnya. Selain itu, pada bagian akhir bab ini akan disajikan saran-saran berdasarkan hasil analisis dan temuan penelitian untuk pengembangan penelitian di masa mendatang.

5.1. Kesimpulan

Berdasarkan penelitian yang sudah dilakukan terkait analisis sentimen ulasan pengguna aplikasi menggunakan algoritma klasifikasi *Logistic Regression* dengan pendekatan penelitian berdasarkan representasi vektornya, dapat disimpulkan hasilnya sebagai berikut:

1. Dalam penerapan analisis sentimen menggunakan algoritma klasifikasi *Logistic Regression* pada ulasan aplikasi LinkedIn, berhasil didapatkan akurasi paling optimal dengan rasio pembagian data 80:20 dan penggunaan model *word embedding Fasttext* dengan ukuran vektor=300, dan epoch=25, serta penggunaan model klasifikasi dengan parameter *learning rate*=0.05 dan jumlah iterasi=1000. Hasil yang didapatkan dengan arsitektur tersebut mencapai **akurasi hingga 95.85%, nilai presisi 96.54%, nilai recall 96.38%, dan nilai F1-score 93.01%**.
2. Penggunaan representasi vektor gabungan antara *TF-IDF* dan *Fasttext* menghasilkan akurasi kinerja model yang lebih unggul dibandingkan hanya menggunakan salah satu metode representasi vektor, tetapi membutuhkan waktu pelatihan yang lebih lama dibandingkan penggunaan masing-masing representasi. Kombinasi antara keduanya berhasil menambah variasi fitur pada dataset yang meningkatkan hasil kinerja model.

5.2. Saran

Berikut adalah beberapa saran yang dapat disampaikan berdasarkan penelitian yang sudah dilakukan guna mendukung hasil pengembangan untuk penelitian selanjutnya :

1. Penyaringan atau analisis lebih lanjut terkait tahapan *pre-processing* untuk pembersihan data mentah khususnya pada tahapan *Stopword Removal* dan *Stemming*, agar perubahan data yang terjadi setelah tahapan tersebut tidak mengubah konteks kalimat dan pengejaan pada data teks yang dapat mempengaruhi konteks dan pemaknaan sentimen.
2. Penambahan jumlah dataset untuk menambah variasi pada karakter dataset ulasan agar model dapat mempelajari data dengan lebih baik dan memberikan hasil pembelajaran yang lebih akurat dalam melakukan prediksi.
3. Percobaan skenario pengujian yang lebih bervariasi dengan mencoba parameter lainnya untuk mendapat hasil yang paling optimal dan sesuai dengan jenis data.
4. Penggunaan jenis dataset yang berbeda untuk mengevaluasi kinerja model klasifikasi dalam mengatasi jenis data yang berbeda.
5. Penyeimbangan dataset antara data ulasan yang berlabel sentimen positif dan negatif untuk mengetahui apakah kinerja model akan lebih optimal dalam memproses data.