

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil analisis dan implementasi yang telah dilakukan, maka dapat disimpulkan bahwa :

1. Proses dalam melakukan analisis sentimen pada *Twitter* terhadap boikot produk Israel menggunakan model BERT adalah diawali dengan melakukan studi literatur dan analisis kebutuhan. Kemudian pengumpulan data pada *Twitter* berdasarkan kata kunci, melakukan *preprocessing text* sehingga menghasilkan *tweet* yang bersih dan siap untuk diproses lebih lanjut. Kemudian melakukan penyaringan data dengan menghapus data duplikasi maupun data dengan bahasa asing. Selanjutnya melakukan pelabelan data otomatis untuk melabeli sentimen dengan VADER. *Exploratory data analysis* (EDA) dilakukan agar dapat memahami lebih dalam data yang dianalisis. Kemudian *word embedding* dilakukan untuk mengubah kata-kata dari format teks ke format numerik agar bisa dibaca oleh pembelajaran mesin. Selanjutnya melakukan pembagian data untuk klasifikasi dan melakukan evaluasi model sehingga didapatkan model yang terbaik. Kemudian melakukan visualisasi dan validasi sistem.
2. Terdapat 2 skenario pembagian data yaitu 80:20 dan 70:30 dalam melakukan *training* model dan terdapat 6 skenario untuk membedakan *hyperparameter* yaitu *epoch*, *batch size*, *learning rate* dan *dropout rate*. Dari hasil percobaan, skenario terbaik didapatkan dari skenario pembagian data 80:20 dengan *dropout rate* 0.1, *epoch* 10, *batch size* 32, dan *learning*

rate sebesar $2e-5$ yang memiliki akurasi yang paling tinggi sebesar 0.847 atau sekitar 85%. Nilai presisi sebesar 0.849 atau 85%, nilai *recall* sebesar 0.848 atau 85% , dan nilai *F1-score* sebesar 0.847 atau 85% dengan waktu *training* 1106 detik. Peningkatan jumlah *epoch* dari 5 ke 10 meningkatkan performa model tetapi waktu komputasi juga meningkat. *Batch size* lebih kecil cenderung menghasilkan performa lebih baik dibandingkan *batch size* lebih besar pada *dropout rate* yang sama. Namun, *batch size* kecil dan *epoch* lebih banyak dapat meningkatkan performa tetapi waktu komputasi lebih lama. Penggunaan *dropout rate* 0.3 dapat mencegah terjadinya kenaikan lonjakan *loss* yang signifikan. Hasil uji validasi sistem dengan skenario terpilih mendapatkan nilai akurasi 76%. Model berkinerja baik dalam memprediksi sentimen positif tetapi kesulitan dalam mengidentifikasi sentimen negatif dengan benar.

3. Proses visualisasi sentimen *tweet* disajikan dalam bentuk *website* menggunakan *library* dari *python* yaitu *streamlit*. Visualisasi dilakukan dengan menampilkan grafik untuk setiap sentimen, grafik untuk tren *tweet* per bulan, grafik *tweet* setiap tanggalnya, grafik untuk kata-kata yang sering muncul seperti “boikot”, “produk”, “israel”, “tidak”, “aksi”, dan “dukung”. Selanjutnya grafik yang menunjukkan panjang *tweet*, dan *wordcloud* untuk setiap sentimen. Terdapat fitur untuk melakukan filter sesuai dengan kebutuhan analisis yang dilakukan. Kemudian dibuat halaman untuk validasi dalam bentuk *input text* dan *upload file csv*.

5.2 Saran

Beberapa saran yang dapat diberikan untuk pengembangan penelitian selanjutnya yaitu sebagai berikut.

1. Penelitian selanjutnya dapat menggunakan skenario klasifikasi model dan proses pengujian yang berbeda.
2. Penelitian selanjutnya dapat menggunakan data *tweet* dengan periode yang lebih baru.
3. Pada penelitian selanjutnya dapat memperluas kapabilitas dengan melakukan normalisasi teks berbahasa inggris dalam *tweet*.
4. Dapat melakukan pelabelan yang dapat mendeteksi kalimat yang mengandung sarkasme.