

BAB V

KESIMPULAN

5.1 Kesimpulan

Berdasarkan hasil analisis dan pengujian perbandingan kinerja *Support Vector Machine* (SVM), *Long Short-Term Memory* (LSTM), dan *IndoBERT* dalam analisis sentimen video *review* produk di TikTok, dapat ditarik beberapa kesimpulan sebagai berikut:

1. Kecenderungan Sentimen Pengguna. Proses pelabelan otomatis menggunakan pendekatan *Lexicon-Based* dengan kamus *InSet* dan pengecekan menggunakan manual mendapatkan 1.670 data komentar yang telah melalui tahap *preprocessing* menunjukkan bahwa sentimen positif merupakan sentimen yang paling dominan, yaitu sebanyak 901 komentar (53,95%), diikuti sentimen netral sebanyak 422 komentar (25,27%), dan sentimen negatif sebanyak 347 komentar (20,78%). Hasil visualisasi *word cloud* menunjukkan bahwa komentar positif banyak berkaitan dengan penilaian positif terhadap produk, komentar netral banyak berkaitan dengan informasi dan pertanyaan mengenai produk, sedangkan komentar negatif berkaitan dengan ketidakpuasan dan pengalaman pengguna terhadap produk.
2. Perbandingan Kinerja Model. Berdasarkan pengujian pada skenario pembagian data 80:20, *IndoBERT* memperoleh *accuracy* sebesar 78,44%, LSTM sebesar 77,84%, dan SVM sebesar 73,35%. Pada nilai *precision*, LSTM memperoleh hasil tertinggi sebesar 80,00%,

sedangkan *IndoBERT* sebesar 78,00% dan SVM sebesar 70,00%. Sementara itu, nilai *recall* dan *F1-score* LSTM dan *IndoBERT* sama-sama sebesar 78,00%, sedangkan SVM memperoleh *recall* sebesar 71,00% dan *F1-score* sebesar 73,00%. Hasil tersebut menunjukkan bahwa *IndoBERT* dan LSTM memiliki kinerja yang lebih baik dibandingkan SVM pada *dataset* penelitian ini, dengan *IndoBERT* unggul pada *accuracy* dan LSTM unggul pada *precision*.

3. Hasil Evaluasi Model. Berdasarkan *confusion matrix* pada 334 data uji, SVM menghasilkan 245 prediksi benar, LSTM sebanyak 260 prediksi benar, dan *IndoBERT* sebanyak 262 prediksi benar. Ketiga model menunjukkan bahwa kelas positif lebih banyak berhasil diklasifikasikan dengan benar, sedangkan kelas negatif relatif lebih sulit diklasifikasikan. Perbedaan performa antar metode dipengaruhi oleh karakteristik *dataset*, tahapan *preprocessing*, distribusi kelas, representasi data, serta konfigurasi model yang digunakan. Berdasarkan keseluruhan hasil evaluasi, *IndoBERT* memperoleh *accuracy* tertinggi sebesar 78,44%, sedangkan LSTM memperoleh *precision* tertinggi sebesar 80,00%, dan keduanya memiliki *recall* serta *F1-score* yang sama sebesar 78,00%.

5.2 Saran

Berdasarkan hasil dan keterbatasan penelitian yang telah dilakukan, beberapa saran yang dapat diberikan untuk penelitian selanjutnya adalah sebagai berikut:

1. Penelitian selanjutnya dapat menggunakan jumlah data yang lebih besar dan lebih beragam agar model memperoleh variasi komentar yang lebih luas dari berbagai produk, video, dan pengguna TikTok. Penambahan variasi data diharapkan dapat membantu model dalam mengenali karakteristik bahasa dan sentimen yang lebih beragam.
2. Penelitian selanjutnya dapat mengembangkan proses analisis sentimen dengan mempertimbangkan perkembangan bahasa yang digunakan pada media sosial, khususnya kata tidak baku, singkatan, slang, dan variasi bahasa pada komentar TikTok. Pengembangan tersebut diharapkan dapat membantu meningkatkan kemampuan model dalam memahami komentar dengan karakteristik bahasa yang informal dan beragam.