BAB V
PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian mengenai klasifikasi komentar pengguna Twitter/X terhadap kegagalan Timnas Indonesia lolos Piala Dunia, penelitian ini berhasil mengklasifikasikan sentimen ke dalam tiga kelas (positif, negatif, netral) melalui tahapan pengumpulan data, preprocessing, ekstraksi fitur TF-IDF, pemodelan, dan evaluasi. Penerapan SVM + TF-IDF memberikan performa yang baik dengan Accuracy 0,7588; Precision 0,7647; Recall 0,7557; F1-Score 0,7584; ROC-AUC 0,9599, serta berdasarkan confusion matrix (626 data uji) diperoleh 475 prediksi benar dan 151 prediksi salah, sehingga model cukup andal dalam mengubah opini publik yang tidak terstruktur menjadi informasi sentimen yang lebih terukur. Selain itu, penelitian ini juga membandingkan tiga metode klasifikasi, yaitu SVM, Multinomial Naive Bayes, dan KNN, dan hasilnya menunjukkan bahwa SVM merupakan metode terbaik pada seluruh metrik evaluasi. Secara persentase, SVM lebih unggul dibanding Naive Bayes sebesar ≈19,95% (Accuracy), 14,60% (Precision), 21,40% (Recall), 21,01% (F1-Score), dan 16,04% (ROC-AUC); sementara dibanding KNN, SVM lebih unggul sebesar ≈8,94% (Accuracy), 10,91% (Precision), 7,99% (Recall), 9,71% (F1-Score), dan 11,75% (ROC-AUC). Keunggulan ini menegaskan bahwa SVM efektif untuk data teks berdimensi tinggi dan sparse seperti TF-IDF. Namun demikian, SVM memiliki beberapa kelemahan, yaitu interpretabilitas yang lebih rendah, sensitif terhadap pemilihan parameter, kebutuhan komputasi yang cenderung lebih besar ketika data meningkat, serta masih berpotensi terjadi kesalahan prediksi pada komentar yang ambigu (misalnya pergeseran kelas negatif/positif ke netral). Dengan demikian, tujuan penelitian telah tercapai, yaitu melakukan klasifikasi sentimen publik di Twitter/X, menerapkan TF-IDF dengan tiga metode pembanding, serta mengevaluasi kinerja model multikelas dengan metrik yang relevan, dengan hasil terbaik diperoleh pada SVM.

5.2 Saran

Untuk pengembangan penelitian selanjutnya, disarankan agar proses pelabelan data ditingkatkan dengan menggabungkan pelabelan otomatis dan validasi manual (human annotation) agar kualitas label lebih konsisten, terutama karena bahasa di Twitter/X sering tidak baku dan berpotensi mengandung ironi atau makna ganda. Selain itu, penelitian berikutnya perlu melakukan optimasi dan tuning parameter SVM secara lebih terstruktur, misalnya melalui grid search atau random search dengan cross-validation, untuk menentukan konfigurasi terbaik seperti nilai C, pemilihan kernel (linear/RBF), parameter gamma (untuk RBF), serta class weight guna mengatasi ketidakseimbangan kelas dan meningkatkan kinerja klasifikasi. Hasil SVM yang sudah dituning kemudian dapat dibandingkan kembali dengan algoritma lain seperti Logistic Regression, Naive Bayes, Random Forest, maupun pendekatan berbasis transformer agar diperoleh model yang paling sesuai dengan karakteristik data.

Selanjutnya, cakupan analisis dapat diperluas tidak hanya pada klasifikasi sentimen, tetapi juga pada analisis yang lebih mendalam seperti analisis temporal sentimen (perubahan sentimen dari waktu ke waktu) dan analisis topik untuk mengetahui tema utama yang dominan dalam opini publik. Penelitian juga dapat mempertimbangkan teknik tambahan seperti deteksi sarkasme untuk meningkatkan pemahaman konteks sentimen pada komentar media sosial. Dari sisi data, disarankan memperluas sumber data ke platform digital lain seperti YouTube, Instagram, atau forum daring, agar hasil penelitian lebih representatif terhadap opini publik Indonesia secara umum. Terakhir, aspek etika penelitian media sosial perlu diperjelas, khususnya terkait anonimisasi identitas pengguna dan kepatuhan pada ketentuan penggunaan data publik, agar penelitian tetap bertanggung jawab secara akademik maupun etis.