

## BAB V KESIMPULAN DAN SARAN

### 5.1. Kesimpulan

Berdasarkan keseluruhan *Proses* penelitian yang telah dilakukan, mulai dari tahap pengumpulan data, pra*Proses* teks, pembentukan fitur menggunakan TF-IDF, hingga pelatihan dan pengujian model menggunakan algoritma *Rocchio* (Nearest Centroid), diperoleh beberapa kesimpulan sebagai berikut:

1. Model *Rocchio* berhasil membedakan komentar ke dalam tiga kategori utama, yaitu *Buzzer Pro*, *Buzzer Kontra*, dan *Non Buzzer*, dengan performa yang cukup stabil pada berbagai skenario pembagian data (90:10, 80:20, 70:30, 60:40, dan 50:50). Hasil pengujian menunjukkan bahwa semakin besar *Proporsi* data latih yang digunakan, semakin baik tingkat akurasi yang dicapai oleh model.
2. Nilai akurasi model pada setiap skenario menunjukkan performa yang kompetitif, dengan hasil berkisar antara 75% hingga 82%. Skenario terbaik diperoleh pada pembagian data 90:10, dengan akurasi tertinggi mencapai 82%, *precision* rata-rata sebesar 0.82, *recall* sebesar 0.82, dan *f1-score* sebesar 0.82. Hal ini menunjukkan bahwa model mampu melakukan *Klasifikasi* dengan baik, khususnya dalam mengenali komentar kategori *Buzzer Pro*.
3. Berdasarkan hasil *classification report*, model menunjukkan performa paling konsisten pada kelas *Buzzer Pro*, dengan nilai *recall* mencapai 0.96–0.97 pada sebagian besar skenario, yang berarti model sangat efektif dalam mengenali komentar yang mendukung pemerintah atau kebijakan. Namun, performa model sedikit menurun pada kelas *Non Buzzer*, di mana *recall* cenderung lebih rendah (sekitar 0.58–0.72), yang menunjukkan bahwa model masih memiliki tantangan dalam membedakan komentar netral dari komentar bermuatan politik.
4. Hasil uji validasi model menggunakan data independen sebanyak 347 data komentar menunjukkan bahwa model memiliki akurasi sebesar 0.90 atau 90%, yang menandakan kemampuan generalisasi model terhadap data baru tergolong sangat baik. Model berhasil memprediksi dengan benar seluruh data *Buzzer Pro* dan *Buzzer Kontra*, serta sebagian besar data *Non Buzzer*. Hal ini menunjukkan bahwa model yang dikembangkan mampu mempertahankan

performa stabil meskipun dihadapkan pada data yang belum pernah dilihat sebelumnya.

5. *Proses* analisis akun *buzzer* mampu memberikan gambaran menyeluruh mengenai kecenderungan komentar setiap pengguna serta pola interaksi pada berbagai channel, sehingga hasilnya dapat menjadi dasar yang kuat dalam memahami distribusi opini dan aktivitas *buzzer* secara keseluruhan.

Secara keseluruhan, algoritma *Rocchio* (Nearest Centroid) terbukti efektif digunakan untuk *Klasifikasi* komentar media sosial dalam konteks identifikasi *buzzer*. Model ini memberikan keseimbangan antara akurasi, kecepatan, dan efisiensi, sehingga dapat diterapkan pada sistem analisis komentar publik yang berskala besar maupun real-time.

## 5.2.Saran

Berdasarkan hasil penelitian dan evaluasi yang telah dilakukan, terdapat beberapa saran yang dapat dijadikan acuan untuk pengembangan penelitian selanjutnya maupun penerapan sistem *Klasifikasi* komentar di masa mendatang, yaitu sebagai berikut:

1. Pengembangan Model dan Optimalisasi Fitur Disarankan agar penelitian selanjutnya melakukan pengembangan model dengan membandingkan algoritma lain seperti Support Vector Machine (SVM), Logistic Regression, atau model berbasis Deep Learning seperti Bidirectional LSTM dan BERT, sehingga dapat diperoleh model dengan tingkat akurasi yang lebih tinggi dan kemampuan generalisasi yang lebih baik. Selain itu, penggunaan teknik feature selection atau feature extraction lanjutan dapat meningkatkan efisiensi komputasi dan mengurangi noise pada data teks.
2. Perluasan dan Peningkatan Kualitas Dataset Dataset yang digunakan dalam penelitian ini masih terbatas dalam jumlah dan konteks komentar. Oleh karena itu, disarankan untuk memperluas cakupan dataset dengan menambahkan komentar dari berbagai platform media sosial, topik politik, maupun isu sosial lainnya. Penambahan data yang lebih beragam dapat membantu model mengenali konteks bahasa yang lebih luas, termasuk variasi gaya bahasa, sarkasme, atau penggunaan bahasa campuran (code-mixing).
3. Penerapan Teknik Text PreProcessing yang Lebih Lanjut *Proses* preProcessing pada penelitian ini dapat dikembangkan lebih lanjut dengan

menambahkan tahapan seperti lemmatization, handling slang words, serta negation detection untuk meningkatkan pemahaman model terhadap makna kalimat secara kontekstual. Tahapan ini diharapkan dapat meningkatkan kemampuan model dalam membedakan antara komentar yang bersifat netral, mendukung, maupun menentang kebijakan.

4. Pengujian dengan Skema Validasi yang Lebih Beragam Untuk memperoleh evaluasi performa model yang lebih menyeluruh, penelitian berikutnya dapat menggunakan skema validasi seperti k-fold cross validation atau stratified sampling. Pendekatan ini akan memberikan gambaran yang lebih akurat terhadap stabilitas model pada berbagai kombinasi data latih dan data uji.
5. Pengembangan Sistem Aplikasi dan Integrasi Real-Time Sistem aplikasi yang telah dibangun dapat dikembangkan lebih lanjut agar mampu melakukan analisis komentar secara real-time dengan integrasi ke API media sosial. Selain itu, hasil *Klasifikasi* dapat disajikan dalam bentuk visualisasi interaktif seperti grafik tren dan sebaran opini, sehingga memudahkan pengguna dalam memahami pola komentar publik terhadap suatu isu atau kebijakan.
6. Sistem analisis perlu dikembangkan dengan visualisasi yang lebih informatif dan metode analisis tambahan agar hasil *Klasifikasi* lebih akurat, mudah diinterpretasikan, serta mampu mendukung kebutuhan analisis lanjutan secara lebih komprehensif dalam analisis akun *buzzer*.