

BAB V

PENUTUP

5.1 Kesimpulan

Penelitian ini berhasil menganalisis sentimen keluhan masyarakat terhadap kualitas jaringan internet di Kabupaten Banggai Kepulauan dengan menggunakan dua model pembelajaran mesin, *XLNet* dan *RoBERTa*. Data yang digunakan terdiri dari 646 sampel komentar yang dikumpulkan dari media sosial WhatsApp dan Google Form Kuesioner. Setelah melalui tahapan preprocessing, yang meliputi pembersihan data, normalisasi kata gaul, dan tokenisasi, serta pelabelan sentimen, model *RoBERTa* yang digunakan adalah "wllwo/indonesian-roberta-base-sentiment-classifier" yang telah dilatih sebelumnya untuk klasifikasi sentimen dalam Bahasa Indonesia. Proses pelabelan sentimen dilakukan menggunakan *TextClassificationPipeline* yang memungkinkan model untuk mengklasifikasikan komentar-komentar yang telah dibersihkan menjadi tiga kategori: Positive, Neutral, dan Negative.

Hasil evaluasi menunjukkan bahwa *RoBERTa* memberikan performa yang lebih baik dibandingkan *XLNet*, dengan akurasi mencapai 87,77% dan *f1-score* 87,75%, sementara *XLNet* hanya mencapai 83,45% untuk akurasi dan 83,37% untuk *f1-score*. Dalam hal distribusi sentimen setelah penanganan ketidakseimbangan data melalui teknik *oversampling*, sentimen Negatif mendominasi dengan 38,5% dari total data, diikuti oleh sentimen Netral dan Positif masing-masing dengan 30,8%. Temuan ini menunjukkan bahwa keluhan masyarakat terhadap kualitas jaringan internet di Kabupaten Banggai Kepulauan cenderung lebih banyak bersifat negatif, terkait dengan masalah seperti lambatnya jaringan atau hilangnya koneksi. Secara keseluruhan, penelitian ini memberikan wawasan penting untuk memperbaiki kebijakan dan pengembangan infrastruktur jaringan di daerah tersebut, serta menambah pemahaman tentang penggunaan teknik pembelajaran mesin dalam analisis sentimen pada Bahasa Indonesia.

5.2 Saran

Beberapa hal yang masih dapat dikerjakan dengan lebih baik dan dikembangkan lebih lanjut dalam penelitian ini adalah sebagai berikut:

1. Dataset yang digunakan terbatas pada dua sumber, yaitu WhatsApp dan Google Form. Pengumpulan data dari platform lain dapat memperkaya dataset dan memberikan hasil yang lebih representatif.
2. Meskipun model RoBERTa memberikan hasil yang baik, eksperimen dengan model lain seperti BERT atau XLNet yang lebih disesuaikan dapat meningkatkan kinerja model.
3. Proses pelabelan sentimen bisa ditingkatkan dengan menggunakan dataset yang lebih besar dan model pelabelan yang lebih akurat untuk mengatasi ketidaktepatan pelabelan pada komentar yang ambigu.
4. Keterbatasan sumber daya, seperti waktu dan kapasitas perangkat keras, menjadi tantangan dalam proses pelatihan model. Penggunaan perangkat keras yang lebih kuat dapat mempercepat proses dan meningkatkan akurasi model.
5. Evaluasi model dengan data yang lebih variatif dan tidak hanya terbatas pada satu wilayah atau topik dapat memberikan gambaran yang lebih komprehensif tentang performa model di dunia nyata.
6. Hasil penelitian ini bisa diterapkan dalam aplikasi nyata, seperti sistem pemantauan keluhan masyarakat atau layanan pelanggan, yang dapat membantu meningkatkan kualitas layanan internet.

Dengan pengembangan lebih lanjut di area-area tersebut, penelitian ini bisa memberikan dampak yang lebih besar dalam meningkatkan kualitas layanan internet dan analisis sentimen.