

BAB V

KESIMPULAN DAN PENELITIAN SELANJUTNYA

5.1 Kesimpulan

Berdasarkan hasil penelitian mengenai Analisis Sentimen Pendapat Konsumen Terhadap Layanan Penyedia Internet IndiHome melalui Twitter dengan Metode *Naïve Bayes Classifier* dan *Laplace Smoothing*, maka dapat diambil beberapa kesimpulan sebagai berikut

1. Penelitian ini berhasil mengimplementasikan metode *Naïve Bayes Classifier* dalam menganalisis sentimen pengguna Twitter terhadap layanan IndiHome dengan pendekatan *SEMMA (Sample, Explore, Modify, Model, Asses)*. Dataset yang digunakan berjumlah 2.000 *tweet* dalam rentang waktu tahun 2021-2024 dan dilakukan preprocessing agar dapat diproses dalam klasifikasi.
2. Hasil klasifikasi menunjukkan bahwa mayoritas *tweet* pengguna memiliki sentimen negatif terhadap layanan IndiHome. Dari total 1.400 data yang digunakan sebagai data uji, sebanyak 811 data diklasifikasikan sebagai sentimen negatif dan 588 data sebagai sentimen positif.
3. Berdasarkan hasil visualisasi *wordcloud* dari dataset yang telah dikumpulkan, ditemukan bahwa kata-kata yang sering muncul adalah “indihome”, “jaringan”, “gangguan”, “lemot”, dan “jelek”. Munculnya kata-kata tersebut mencerminkan persepsi negatif yang dominan dari pengguna Twitter terhadap layanan IndiHome. Hal ini menunjukkan bahwa sebagian besar pengguna merasa tidak puas dengan layanan yang diberikan oleh IndiHome. Temuan ini sejalan dengan hasil klasifikasi sentimen yang juga menunjukkan dominasi opini negatif dalam data yang dianalisis.
4. Model *Naïve Bayes* memberikan performa yang cukup baik dalam memprediksi sentimen, sesuai dengan evaluasi menggunakan metode *Split*

Validation dengan hasil akurasi sebesar 70,12%. Namun, hasil evaluasi juga mengindikasikan bahwa model mengalami kesulitan dalam mengenali sentimen positif, dengan nilai *precision* hanya sebesar 47,5% meskipun pada *recall* mencapai 70,69%.

5. Penambahan teknik *Laplace Smoothing* pada metode *Naive Bayes* terbukti memberikan peningkatan performa model dalam proses klasifikasi. Berdasarkan pengujian terhadap 10 data uji yang sama, metode *Naive Bayes* tanpa *Laplace Smoothing* hanya berhasil mengklasifikasikan 8 data dengan benar atau dengan akurasi sebesar 80%, sedangkan 2 data positif salah terklasifikasi sebagai negatif. Setelah diterapkan *Laplace Smoothing*, seluruh kata pada data uji memperoleh probabilitas lebih dari nol pada setiap kelas sehingga semua data uji dapat terklasifikasi dengan benar hingga mendapatkan akurasi sebesar 100%. Hal ini menunjukkan bahwa *Laplace Smoothing* mampu mengatasi permasalahan probabilitas nol yang muncul akibat ketidakhadiran kata-kata tertentu dalam data latih, sehingga model menjadi lebih adaptif dan akurat dalam menangani data baru yang bervariasi.
6. Berdasarkan hasil *Area Under Curve (AUC)*, performa model dalam skenario normal menghasilkan nilai 0,594 yang mengindikasikan performa sedang. Hal ini menunjukkan bahwa model masih memiliki ruang untuk ditingkatkan, terutama dalam menangani distribusi data yang tidak seimbang.

5.2 Penelitian Selanjutnya

Adapun saran yang dapat diberikan untuk pengembangan penelitian selanjutnya adalah sebagai berikut :

1. Menggunakan data set yang lebih besar dan menyeimbangkan jumlah data positif dan negatif agar model tidak terlalu condong dalam memprediksi satu kelas tertentu dan dapat mengenali pola sentimen secara lebih adil.
2. Menggunakan algoritma pembandingan, seperti *Support Vector Machine (SVM)*, *Random Forest*, atau algoritma berbasis *deep learning* guna

mengevaluasi dan membandingkan performa klasifikasi dengan algoritma lain.

3. Meningkatkan teknik *feature extraction* atau pembobotan kata, misalnya dengan menggunakan *word embedding* seperti *Word2Vec* atau *FastText* agar representasi kata lebih kontekstual dibandingkan hanya dengan menggunakan TF-IDF.
4. Memperluas klasifikasi menjadi tiga kelas, yaitu positif, negatif, dan netral.
5. Menggabungkan data dari berbagai platform media sosial agar dapat memberikan pandangan yang lebih luas terhadap persepsi publik secara lintas media.
6. Melakukan analisis sentimen berbasis waktu atau *temporal analysis* agar dapat melihat perubahan persepsi publik terhadap suatu produk atau layanan dari waktu ke waktu.