

BAB 5 SIMPULAN DAN SARAN

5.1 Simpulan

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa:

1. Penelitian ini berhasil mengimplementasikan algoritma *Support Vector Machine (SVM)* untuk mengklasifikasikan opini masyarakat dalam tiga kategori sentimen, yaitu positif, netral, dan negatif. Proses klasifikasi diawali dengan tahap *preprocessing* data teks, dilanjutkan dengan transformasi fitur menggunakan *TF-IDF* dengan rentang *N-gram* (1,2) dan penghapusan kata jarang dengan *min_df=3*. Untuk mengatasi ketidakseimbangan jumlah kelas, digunakan teknik *SMOTE* sebagai metode oversampling. Model akhir menggunakan *Linear SVM* dengan parameter $C = 1.0$ yang ditentukan berdasarkan evaluasi menggunakan *validation curve*. Model ini tidak hanya menghasilkan prediksi, tetapi juga mampu menginterpretasikan fitur-fitur yang paling mempengaruhi klasifikasi, menjadikannya transparan dan layak untuk diterapkan dalam konteks nyata.
2. Evaluasi model dilakukan menggunakan metrik akurasi, precision, recall, dan *F1-score* yang dirinci melalui *classification report* dan *confusion matrix*. Hasil evaluasi terhadap data uji menunjukkan bahwa model memiliki akurasi sebesar 78,73%, dengan nilai *macro average F1-score* sebesar 0,7614. Model menunjukkan performa terbaik dalam mengenali opini negatif (*F1-score* 0,8646), disusul oleh positif (0,7879), dan netral (0,6316). Berdasarkan hasil ini, dapat disimpulkan bahwa model dapat secara efektif mengukur dan mengelompokkan sentimen masyarakat terhadap insiden peretasan Indodax secara sistematis dan dapat diandalkan.

5.2 Saran

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

1. Pada penelitian selanjutnya disarankan untuk menerapkan pendekatan tanpa penghapusan *stopword*, terutama jika menggunakan metode *lexicon* seperti

InSet, karena penghapusan *stopword* berpotensi menghilangkan kata penting seperti tidak atau kurang yang memengaruhi polaritas sentimen. Pendekatan ini juga dapat dibandingkan dengan penggunaan *stopword* kustom untuk mengetahui dampaknya terhadap hasil analisis.

2. Untuk meningkatkan performa klasifikasi terutama pada kelas netral, disarankan untuk mengeksplorasi metode representasi teks lain seperti *word2vec*, *GloVe*, atau model berbasis konteks seperti *BERT*.
3. Perbandingan performa antar algoritma belum dilakukan dalam penelitian ini. Oleh karena itu, disarankan untuk membandingkan model *SVM* dengan algoritma klasifikasi lainnya seperti *Random Forest*, *XGBoost*, atau *deep learning*.
4. Implementasi antarmuka visual atau dashboard yang menampilkan hasil klasifikasi dan fitur dominan secara interaktif dapat menjadi kontribusi praktis yang bernilai bagi pengguna non-teknis.
5. Penelitian ini berfokus pada satu kasus peristiwa. Untuk meningkatkan generalisasi model, perlu dilakukan pengujian terhadap data opini publik dari isu-isu lain yang relevan, khususnya dalam ranah keamanan digital atau keuangan.

