

BAB V PENUTUP

5.1 Kesimpulan

Berdasarkan rumusan masalah yang telah ditetapkan serta hasil pengujian dan pembahasan pada Bab IV, dapat disimpulkan bahwa model Naive Bayes dan Logistic Regression sama-sama mampu melakukan deteksi berita hoaks berbahasa Indonesia dengan cukup baik ketika direpresentasikan menggunakan fitur TF-IDF. Tahapan preprocessing yang meliputi *cleaning*, *tokenizing*, *case folding*, dan *stemming* berperan penting dalam meningkatkan kualitas data teks sehingga representasi TF-IDF yang dihasilkan menjadi lebih efektif dalam menangkap karakteristik kata yang membedakan berita hoaks dan berita non-hoaks. Hasil evaluasi menunjukkan bahwa kedua model dapat mengklasifikasikan berita hoaks dengan tingkat akurasi yang tinggi, menandakan bahwa pendekatan klasifikasi berbasis TF-IDF relevan untuk permasalahan deteksi berita hoaks.

Berdasarkan perbandingan kinerja menggunakan metrik evaluasi berupa akurasi, presisi, recall, dan F1-score, model Logistic Regression menunjukkan performa yang cenderung lebih optimal dibandingkan dengan model Naive Bayes pada dataset yang digunakan dalam penelitian ini. Logistic Regression mampu memanfaatkan bobot fitur TF-IDF secara lebih stabil sehingga menghasilkan keseimbangan yang lebih baik antara kemampuan mengenali berita hoaks dan menghindari kesalahan klasifikasi. Meskipun demikian, Naive Bayes tetap menunjukkan performa yang kompetitif dengan keunggulan pada kesederhanaan model dan efisiensi komputasi. Dengan demikian, dapat disimpulkan bahwa Logistic Regression merupakan model yang paling optimal untuk deteksi berita hoaks berbahasa Indonesia berdasarkan metrik evaluasi yang digunakan, sementara Naive Bayes tetap layak digunakan sebagai alternatif yang ringan dan efisien.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran yang dapat dijadikan pertimbangan untuk penelitian selanjutnya, antara lain:

1. Penelitian selanjutnya disarankan untuk menggunakan dataset dengan jumlah data yang lebih besar dan sumber yang lebih beragam agar model yang dihasilkan memiliki kemampuan generalisasi yang lebih baik.
2. Dapat dilakukan perbandingan dengan algoritma klasifikasi lainnya, seperti Support Vector Machine (SVM), Random Forest, atau metode deep learning untuk mengetahui performa yang lebih optimal dalam mendeteksi berita hoaks.
3. Penelitian selanjutnya dapat mengembangkan metode representasi fitur lain, seperti Word2Vec, FastText, atau BERT, guna membandingkan efektivitasnya dengan metode TF-IDF.
4. Sistem deteksi berita hoaks yang dihasilkan pada penelitian ini dapat dikembangkan lebih lanjut menjadi aplikasi atau sistem pendukung keputusan yang dapat membantu masyarakat dalam memverifikasi informasi.

