

BAB V

SIMPULAN DAN SARAN

5.1 Simpulan

Pengembangan model *machine learning* untuk mengidentifikasi berita *hoax* merupakan langkah yang krusial dalam menanggapi permasalahan penyebaran informasi palsu di era digital. Dalam pengembangan skripsi, metodologi yang dapat diadopsi mencakup pengumpulan data berita dari berbagai sumber, *preprocessing* teks, ekstraksi fitur menggunakan *tf-idf*, dan penerapan algoritma klasifikasi seperti *Logistic Regression*, *Decision Tree*, *Gradient Boosting*, dan *Random Forest*.

Proses ini melibatkan tahapan pengumpulan dataset yang mencakup berita valid dan *hoax*, pelabelan data, serta langkah-langkah *prapemrosesan* teks untuk meningkatkan kualitas data. Analisis data eksploratif menjadi langkah awal dalam memahami karakteristik dataset. Ekstraksi fitur menggunakan *tf-idf* menjadi langkah penting untuk mengubah teks menjadi representasi numerik yang dapat diproses oleh algoritma *machine learning*.

Penerapan algoritma klasifikasi seperti *Logistic Regression*, *Decision Tree*, *Gradient Boosting*, dan *Random Forest* merupakan langkah-langkah utama dalam melatih model dan melakukan evaluasi kinerja. Evaluasi menggunakan metrik standar seperti akurasi, presisi, *recall*, dan *f1-score* memberikan gambaran holistik tentang keefektifan model dalam mengidentifikasi berita *hoax*.

Analisis hasil bertujuan untuk memberikan pemahaman mendalam tentang kinerja model, mengidentifikasi potensi kesalahan, dan menganalisis fitur-fitur yang paling berpengaruh dalam klasifikasi berita *hoax*. Keseluruhan proses ini diharapkan dapat memberikan kontribusi dalam mengembangkan solusi cerdas untuk memerangi penyebaran berita *hoax* dan membangun lingkungan informasi yang lebih dapat dipercaya di era digital.

5.2 Saran

Berikut beberapa saran untuk pengembangan lebih lanjut:

1. Perluasan Analisis Kesalahan: Selain mengidentifikasi kesalahan prediksi model, perlu melakukan analisis lebih mendalam terhadap kesalahan tersebut. Fokus pada karakteristik linguistik atau konten berita yang sering menyebabkan kesalahan klasifikasi. Hal ini dapat membantu memperbaiki model dan meningkatkan akurasi.
2. Pertimbangkan Penggunaan Teknik NLP Lanjutan: Teknik-teknik *Natural Language Processing* (NLP) lanjutan seperti *Word Embeddings* atau model bahasa yang telah dilatih sebelumnya mungkin dapat meningkatkan pemahaman terhadap konten berita. Eksplorasi teknik-teknik ini dapat memberikan wawasan tambahan dalam analisis berita.
3. Perbandingan Dengan Model Lain: Selain algoritma klasifikasi yang telah digunakan, pertimbangkan untuk membandingkan performa model dengan menggunakan arsitektur atau algoritma lain. Ini dapat memberikan pemahaman lebih baik tentang keunggulan dan kelemahan model yang

dikembangkan.

4. Implementasi pada *Dataset* yang Berbeda: Jika memungkinkan, uji model pada dataset yang berbeda untuk memeriksa sejauh mana model dapat menggeneralisasi ke berbagai konteks. Hal ini penting untuk menilai keberlanjutan dan keandalan model di luar *dataset* pelatihan.
5. Sesuaikan dengan Kebutuhan Pemangku Kepentingan: Pertimbangkan *feedback* dari pemangku kepentingan terkait dengan kebutuhan dan harapan mereka terhadap model. Sesuaikan pengembangan model dengan kebutuhan praktis yang spesifik.
6. Penelitian Masa Depan: Dalam bagian Saran, tambahkan wawasan mengenai arah penelitian masa depan. Misalnya, pengembangan model yang lebih kompleks, integrasi dengan sumber data sosial media, atau adaptasi model terhadap tren berita baru.

Dengan penerapan saran ini, penelitian di masa depan dapat memperkuat deteksi berita *hoax* dan membantu mengurangi dampak penyebaran informasi palsu.