

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dari hasil penelitian ini yang didapatkan berdasarkan analisis, perancangan sistem, perbandingan kedua *model machine learning* yang digunakan, dan implementasi sistem dua model *machine learning* kedalam aplikasi maka kesimpulan yang didapatkan adalah:

1. Hasil dari implementasi *machine learning* dengan algoritma *Support Vector Machine* dan *Random Forest* dalam mengembangkan model yang dapat melakukan klasifikasi konten negatif pada website berbahasa indonesia dapat dilakukan.
2. Model *machine learning* untuk *Support Vector Machine* memiliki metrik yang lebih baik yaitu 97% *accuracy*, 90% *precision*, 91% *recall*, dan 90% *F1-score*. sedangkan *Random Forest* hasil evaluasi modelnya adalah 92% *accuracy*, 77% *precision*, 86% *recall*, 76% *F1-score*. Dalam kasus mendeteksi teks konten negatif model svm bekerja lebih baik dibandingkan rf.

5.2 Saran

Dari hasil penelitian ini sampai dengan implementasi *machine learning* ada saran yang dapat digunakan untuk mengembangkan sistem kedepannya:

1. Menyeimbangkan data atau memperbanyak data untuk melatih model *machine learning* yang digunakan. Dalam penelitian ini data yang didapatkan masih *imbalance* sehingga hasil klasifikasi untuk penipuan memiliki hasil yang lebih rendah dibandingkan dengan kelas lain.
2. Untuk membersihkan *stopwords* yang ada pada data teks dengan lebih baik dapat menambahkan kata-kata yang dianggap sebagai *stopwords* kedalam *list* dan memasukkan secara manual.
3. Dapat menggunakan model *deep learning* terbaru yang dikhususkan untuk teks seperti *BERT* atau *Transformer*. Karena penggunaan model tersebut masih sedikit dalam imlementasi data yang berbahasa Indonesia.

4. Pengembangan sistem kedepannya mendeteksi konten negatif dapat menambahkan kelas yang akan membantu dalam mengklasifikasi berita hoaks.
5. Untuk *machine learning* yang digunakan dapat dilakukan *hyperparameter tuning* untuk mengoptimasi dan mendapatkan hasil yang lebih baik dengan *Grid Search* yang berfungsi untuk mencari parameter terbaik untuk model.

