

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil analisis yang telah dipaparkan pada Bab IV, penelitian ini berhasil menjawab rumusan masalah mengenai perbandingan kinerja lima algoritma machine learning *Naive Bayes*, *Support Vector Machine (SVM)*, *Random Forest*, *K-Nearest Neighbor (k-NN)*, dan *Logistic Regression* dalam mendeteksi promosi perjudian online pada komentar YouTube. Evaluasi dilakukan dengan menggunakan metrik akurasi, *precision*, *recall*, *F1-score*, serta *AUC* untuk memperkuat analisis performa. Dari hasil pengujian yang dilakukan, dapat disimpulkan beberapa hal sebagai berikut:

1. Penelitian ini berhasil menjawab rumusan masalah terkait perbandingan kinerja *Naive Bayes*, *Support Vector Machine (SVM)*, *Random Forest*, *K-Nearest Neighbor (k-NN)*, dan *Logistic Regression*. Hasil pengujian menunjukkan bahwa seluruh algoritma mencapai akurasi di atas 0,90 pada dataset berbahasa Indonesia yang direpresentasikan menggunakan *Term Frequency-Inverse Document Frequency (TF-IDF)* tanpa pembersihan data lanjutan.
2. *Random Forest* terbukti menjadi algoritma dengan kinerja terbaik. Model ini memperoleh akurasi 0,97, *F1-score* mendekati sempurna, serta *AUC* 0,99. Keunggulan ini disebabkan oleh kemampuannya menangkap interaksi kompleks antar fitur *TF-IDF* serta ketahanannya terhadap variasi penulisan promosi yang tidak terstruktur. Dengan demikian, *Random Forest* direkomendasikan sebagai model utama dalam pengembangan sistem moderasi otomatis pada platform YouTube.
3. *Support Vector Machine (SVM)* menempati posisi kedua dengan performa yang sangat kompetitif, khususnya pada kemampuan mendeteksi pola bahasa non-linier dan penggunaan simbol atau slang. Dengan *recall* 0,94 pada komentar promosi, *SVM* menunjukkan efektivitas tinggi dalam klasifikasi konten berbahaya.

4. *Naive Bayes* menunjukkan keunggulan pada metrik *recall* (0,98), yang berarti model ini sangat efektif dalam mengenali kata kunci yang sering digunakan dalam promosi judi online. Namun, *precision* yang relatif rendah menunjukkan adanya kecenderungan menghasilkan *false positive*, sehingga penggunaannya memerlukan penyesuaian untuk mengurangi kesalahan klasifikasi pada komentar non-promosi.
5. *Logistic Regression* memiliki akurasi 0,96 dan *AUC* sempurna (1.00), yang menunjukkan kemampuan pemisahan kelas dan peringkat probabilitas yang sangat baik. Hal ini menandakan bahwa model tidak hanya mampu membedakan kelas dengan presisi tinggi, tetapi juga memiliki tingkat keyakinan prediksi yang andal, sehingga layak dipertimbangkan untuk implementasi nyata.
6. *K-Nearest Neighbor (k-NN)* memperoleh kinerja terendah dengan akurasi 0,92 dan *recall* 0,89. Rendahnya *recall* menunjukkan model ini sering gagal mendeteksi komentar promosi yang tersamar, akibat pengaruh *curse of dimensionality* dari representasi *TF-IDF*.
7. Hasil penelitian ini membuktikan bahwa metode representasi *TF-IDF*, meskipun tanpa proses pembersihan data seperti *stopword removal* atau *stemming*, tetap mampu menangkap informasi linguistik yang relevan untuk klasifikasi komentar berbahasa Indonesia yang penuh variasi penulisan. Hal ini penting karena pada kasus deteksi promosi judi online, proses pembersihan data yang berlebihan justru berisiko menghilangkan kata atau simbol penting yang menjadi indikator promosi.

Dengan mempertimbangkan tujuan penelitian untuk meminimalkan lolosnya komentar promosi judi online, *recall* menjadi metrik utama yang menentukan efektivitas model. Oleh karena itu, *Random Forest* dinilai sebagai algoritma yang paling sesuai untuk diterapkan pada sistem deteksi komentar promosi judi online, diikuti oleh *SVM* sebagai alternatif kedua yang juga menjanjikan.

5.2 Saran

Berdasarkan hasil penelitian yang telah diperoleh, terdapat beberapa saran yang dapat dipertimbangkan untuk pengembangan penelitian selanjutnya:

1. Mengingat penelitian ini menunjukkan *TF-IDF* efektif bahkan tanpa pembersihan data, pengembangan ke depan dapat mengeksplorasi variasi parameter *TF-IDF* (misalnya pengaturan *ngram_range*, *min_df*, dan *max_df*) serta tokenisasi untuk mengoptimalkan hasil, khususnya dalam menangkap frasa multi-kata seperti "slot gacor" atau "bonus harian".
2. Melakukan eksperimen dengan kombinasi beberapa model kuat (misalnya *Random Forest* + *SVM*) untuk memanfaatkan keunggulan masing-masing algoritma, tanpa bergantung pada *embedding* yang membutuhkan teks bersih.
3. Melakukan optimasi *hyperparameter* secara lebih mendalam untuk setiap algoritma, misalnya jumlah *estimator* pada *Random Forest* atau nilai *C* dan *gamma* pada *SVM*, sehingga kinerja model dapat dimaksimalkan dengan tetap menggunakan representasi *TF-IDF*.
4. Menjadikan hasil penelitian ini sebagai baseline untuk pengembangan sistem moderasi komentar otomatis pada platform digital, serta melakukan pengujian lanjutan menggunakan data *real-time* dan pembaruan model secara periodik agar sistem tetap adaptif terhadap perkembangan pola bahasa dan strategi penyamaran promosi judi online.

Melalui penerapan saran-saran tersebut, diharapkan penelitian selanjutnya dapat menghasilkan model deteksi yang lebih akurat, adaptif, dan mampu menghadapi tantangan baru dalam identifikasi konten berbahaya di berbagai platform digital.