BAB 5
KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, dapat diambil beberapa poin kesimpulan sebagai berikut:

1. Sistem deteksi *phising* berbasis URL berhasil dikembangkan dengan menggunakan algoritma *Decision Tree* (CART) dan fitur-fitur struktural dari URL.
2. Data yang digunakan berasal dari dua sumber utama:
    a. URL *phising* sebanyak 54.647 data dari *PhishTank* (setelah dibersihkan menjadi 54.630),
    b. URL *legitimate* sebanyak 54.000 dari *Tranco,* dengan total gabungan 108.630 data.
3. Proses ekstraksi fitur menghasilkan 11 fitur penting yang mencerminkan karakteristik URL *phising* dan *legitimate*, seperti panjang URL, jumlah titik, keberadaan HTTPS, dan jumlah karakter khusus.
4. Model *Decision Tree* CART menunjukkan performa sangat baik, dengan:
    a. Akurasi testing: 99,49%
    b. Akurasi training: 99,57%
    c. Rata-rata akurasi cross-validation (5-fold): 99,56%
    d. Precision, recall, dan f1-score masing-masing 0,99 untuk kedua kelas (*phising* dan *legitimate*).
5. Visualisasi pohon keputusan menunjukkan bahwa fitur seperti *url_length, num_at, num_slashes, num_digits, num_special_chars, has_https, num_subdomains* menjadi faktor paling berpengaruh dalam klasifikasi.
6. Evaluasi menggunakan *Confusion Matrix,* ROC *Curve*, dan *Precision-Recall Curve* menunjukkan bahwa model mampu meminimalisasi kesalahan klasifikasi dan bekerja baik dalam skenario ketidakseimbangan data.

7. Model tidak mengalami *overfitting*, dibuktikan dengan perbedaan akurasi yang sangat kecil antara data latih dan uji serta hasil *cross-validation* yang stabil.
8. Uji prediksi *real-time* terhadap 10 URL baru menunjukkan kemampuan generalisasi model yang baik terhadap data di luar pelatihan, dengan hasil yang akurat dan konsisten.

5.2. Saran

Dari hasil penelitian yang telah diperoleh dari pembahasan beberapa saran yang dapat diberikan adalah:

1. Perluasan fitur deteksi, dengan memasukkan informasi tambahan seperti:
    a. Metadata WHOIS,
    b. Konten HTML,
    c. Analisis JavaScript.
2. Pembaruan dataset secara berkala, agar sistem tetap relevan dengan pola serangan *phising* terbaru yang terus berkembang.
3. Implementasi sistem dalam bentuk aplikasi nyata, seperti aplikasi desktop atau *web-based* untuk mendukung pemanfaatan secara langsung oleh pengguna umum.
4. Eksplorasi algoritma *machine learning* lain, seperti *Random Forest* atau *ensemble learning*, untuk meningkatkan performa dan ketahanan model terhadap variasi data.
5. Kolaborasi dengan penyedia layanan keamanan siber, untuk integrasi sistem ke dalam platform deteksi *phising* yang digunakan secara luas