BAB V

SIMPULAN DAN SARAN

5.1 Simpulan

Penelitian ini diawali dengan permasalahan tingginya ancaman phishing yang menargetkan sektor perbankan, termasuk PT Bank Central Asia Tbk (BCA), yang mengakibatkan potensi kerugian finansial dan reputasi. Dengan semakin canggihnya teknik serangan, seperti penyamaran melalui URL palsu yang menyerupai domain resmi bank, nasabah dan karyawan perusahaan menjadi rentan terhadap penipuan digital. Untuk itu, penelitian ini bertujuan mengembangkan model deteksi URL phishing berbasis machine learning yang efektif dan implementasinya dalam sistem berbasis web.

Penelitian ini berhasil mengembangkan model deteksi URL phishing untuk PT Bank Central Asia Tbk dengan menggunakan empat algoritma machine learning, yaitu Support Vector Machine (SVM), Decision Tree, Naïve Bayes, dan Random Forest. Model ini dibuat melalui pendekatan CRISP-DM yang meliputi pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, dan implementasi. Dari hasil pengujian, algoritma SVM menunjukkan kinerja terbaik dengan Akurasi pelatihan adalah 99.82%, dan akurasi pengujian adalah 99.88%, diikuti oleh Decision Tree Akurasi pelatihan (Training Accuracy) adalah 99.26%, dan akurasi pengujian (Testing Accuracy) adalah 98.77%, Random Forest Akurasi pelatihan adalah 99.08%, dan akurasi pengujian adalah 98.65%, dan Naïve Bayes Akurasi pelatihan adalah 92.99%, dan akurasi pengujian adalah 92.51%. Model terbaik diimplementasikan pada sistem berbasis web untuk mendeteksi URL phishing yang relevan dengan data dari PT Bank Central Asia Tbk, sehingga sistem ini dapat mendukung mitigasi ancaman phishing secara lebih efektif.

Evaluasi komparatif terhadap kinerja algoritma menunjukkan bahwa SVM unggul dalam mendeteksi URL phishing dengan tingkat akurasi yang lebih

tinggi dibandingkan model lainnya. Keunggulan ini didukung oleh pemrosesan data yang lebih presisi dan kemampuan algoritma dalam menangani pola data yang kompleks. Meskipun algoritma lain seperti Random Forest dan Decision Tree juga memberikan performa yang baik, hasil menunjukkan bahwa mereka tidak dapat menyaingi SVM dalam hal akurasi dan generalisasi pada data pengujian. Naïve Bayes, meskipun akurat untuk klasifikasi sederhana, menunjukkan keterbatasan ketika dihadapkan pada data yang lebih kompleks.

Tahapan pengembangan sistem deteksi berbasis web dilakukan dengan mengintegrasikan model terbaik ke dalam antarmuka yang dirancang untuk mempermudah pengguna dalam memvalidasi URL yang mencurigakan. Sistem ini tidak hanya memberikan hasil yang cepat dan akurat tetapi juga dirancang untuk mudah digunakan oleh karyawan dan nasabah PT Bank Central Asia, Tbk . Efektivitas sistem diuji melalui pengujian langsung dan umpan balik dari pengguna, yang menunjukkan bahwa sistem ini mampu membantu dalam mengidentifikasi URL phishing dengan akurasi tinggi, sehingga memberikan perlindungan tambahan terhadap serangan siber yang semakin canggih. Implementasi sistem ini diharapkan dapat meningkatkan kepercayaan nasabah terhadap layanan digital PT Bank Central Asia, Tbk serta memperkuat strategi keamanan perusahaan.

5.2 Saran

Untuk pengembangan lebih lanjut agar hasil penelitian ini dapat lebih optimal, ada beberapa saran yang dapat dipertimbangkan:

1. Memperluas dan Meningkatkan Kualitas Dataset untuk Pelatihan Model
   Dataset yang lebih luas dapat membantu meningkatkan akurasi model dalam mendeteksi URL phishing. Semakin banyak data yang digunakan, semakin baik model dapat mengenali pola-pola yang umum maupun yang jarang ditemukan dalam URL phishing, sehingga meningkatkan kemampuan model untuk melakukan generalisasi dan

memberikan prediksi yang akurat pada data baru. Dengan dataset yang lebih beragam, model dapat mengenali berbagai variasi dari URL yang mencurigakan, sehingga dapat meminimalisir risiko kesalahan deteksi.

2. Menambah Jumlah URL dalam Dataset untuk Memperluas Lingkup Deteksi Penambahan jumlah URL dalam dataset tidak hanya meningkatkan akurasi model, tetapi juga memperluas cakupan deteksi. Dengan semakin banyaknya URL yang tersedia, model dapat mengidentifikasi pola yang berbeda pada URL dari berbagai sumber dan tipe, sehingga akan lebih tangguh dalam menghadapi URL baru yang mungkin berbeda dari yang telah dilatihkan. Hal ini juga penting untuk mengurangi bias data dan memastikan bahwa model tidak hanya berfungsi baik pada dataset tertentu saja, tetapi mampu mendeteksi phishing pada berbagai jenis URL secara efektif.
3. Mempertimbangkan Penggunaan Model dan Metode Evaluasi yang Berbeda, seperti Gradient Boosting atau XGBoost Model lain yang lebih canggih, seperti Gradient Boosting atau XGBoost, dapat dipertimbangkan untuk memberikan hasil yang lebih akurat. Kedua model ini merupakan metode ensemble yang menggabungkan beberapa model sederhana untuk membentuk model yang kuat dan mampu menangani data yang kompleks. XGBoost, sebagai varian dari Gradient Boosting, dikenal dengan kemampuannya dalam menangani data besar dengan waktu pelatihan yang relatif cepat serta akurasi yang tinggi. Selain itu, metode evaluasi yang tepat juga harus dipertimbangkan untuk memastikan model benar-benar efektif dan tidak overfitting.