

BAB V

SIMPULAN DAN SARAN

5.1 Simpulan

Hasil dari penelitian ini menerapkan model algoritma klasifikasi menggunakan *Random Forest*, *Support Vector Machine*, dan *K-Nearest Neighbors* untuk memprediksi *URL web phishing* dengan kasus yang semakin meningkat setiap tahunnya. Penelitian ini menerapkan metode analisis menggunakan kerangka kerja *CRISP-DM*, meliputi pemahaman bisnis, pemahaman data, persiapan data, pemodelan, dan evaluasi.

1. Tahap pemahaman bisnis sangat penting untuk memahami tujuan penelitian dari sudut pandang bisnis, yang kemudian disesuaikan menjadi definisi permasalahan data mining. Tahap pemahaman data digunakan untuk membantu dalam mengidentifikasi data yang mendukung penelitian ini, kemudian diolah pada tahap persiapan data. Data tersebut digunakan untuk membuat grafik, menghitung nilai *confusion matrix*, serta menentukan model terbaik yang cocok untuk permasalahan deteksi *URL website phishing*.
2. Evaluasi dilakukan untuk menganalisis hasil perhitungan nilai *confusion matrix* dan memeriksa model terbaik yang akan dipilih untuk membantu dalam mendukung pengambilan keputusan.
3. Penelitian ini menggunakan algoritma *Random Forest*, *K-Nearest Neighbors*, dan *Support vector Machine* dengan memeriksa nilai *accuracy*, *precision*, *f1 score*, and *recall*.
4. Hasil dari evaluasi menggunakan 30 atribut berdasarkan hasil komparasi yang telah dilakukan bahwa algoritma *Random Forest* menjadi model pilihan terbaik dalam menganalisis *URL web phishing*. Dengan nilai akurasi sebesar 96.7% dan *f1 score* sebesar 97.1%.
5. Hasil dari evaluasi menggunakan 5 atribut terpenting berdasarkan hasil komparasi yang telah dilakukan bahwa algoritma *Random Forest* menjadi

model pilihan terbaik dalam menganalisis *URL web phishing*. Dengan nilai akurasi sebesar 92.8% dan *f1 score* sebesar 93.6%.

5.2 Saran

Hasil dari studi ini menyarankan beberapa rekomendasi untuk penelitian lanjutan dengan fokus pada topik yang serupa di masa mendatang. Rekomendasi yang disarankan meliputi:

6. Karena adanya keterbatasannya waktu dalam penelitian ini, maka penelitian ini dilakukan dalam jangkauan tahapan evaluasi. Saran untuk peneliti selanjutnya dapat menggunakan hasil studi ini sebagai rujukan untuk melakukan penelitian sejenis sampai tahap deployment atau implementasi, sehingga produk nyatanya dapat digunakan untuk menangani kasus *URL website phishing* yang bermanfaat untuk masyarakat.
7. Penelitian ini dilakukan dengan menggunakan model algoritma klasifikasi yaitu *Random Forest*, *SVM*, dan *KNN*. Untuk penelitian selanjutnya dapat dilakukan optimasi dengan cara mengkomparasikan model yang telah dibuat dengan menggunakan algoritma *clustering* agar dapat menghasilkan *output* yang berbeda, sehingga nantinya dapat digunakan untuk menjadi perbandingan terhadap model klasifikasi yang dapat mengetahui model algoritma terbaik dalam mendeteksi *URL website phishing* yang ada.

