BAB 1
PENDAHULUAN

1.1 Latar Belakang

Internet dan media digital kini telah menjadi bagian yang tidak terpisahkan dari kehidupan manusia. Situs web, sebagai salah satu media digital yang populer, digunakan untuk berbagai aktivitas, seperti berbelanja *online*, mencari informasi, hingga berinteraksi secara sosial. Kemudahan akses dan efisiensi waktu menjadikan penggunaan situs web semakin meluas. Didukung oleh inovasi teknologi yang pesat, jumlah situs web pun terus bertambah sehingga semakin mempererat hubungan antara internet dan kehidupan sehari-hari.

Namun, di balik berbagai kemudahan tersebut, terdapat ancaman serius terhadap keamanan pengguna, salah satunya adalah serangan *phishing*. *Phishing* merupakan upaya penipuan melalui situs web palsu yang dibuat menyerupai situs resmi dengan tujuan mencuri data sensitif, seperti kata sandi, nomor kartu kredit, hingga identitas pribadi. Serangan ini tidak hanya merugikan individu, tetapi juga dapat mengancam keamanan data institusi dan perusahaan.

Di Indonesia, serangan *phishing* menunjukkan tren peningkatan yang mengkhawatirkan. Berdasarkan laporan Badan Siber dan Sandi Negara (BSSN) periode 2023–2024, tercatat 26,7 juta aktivitas *phishing* sepanjang tahun 2024, meningkat drastis dibandingkan tahun sebelumnya. Selain itu, sebanyak 56,1 juta data terekspos di *darknet*, yang mengancam setidaknya 461 institusi. Target serangan tidak hanya mencakup sektor finansial, tetapi juga infrastruktur vital dan layanan publik dengan memanfaatkan teknik rekayasa sosial serta eksploitasi kerentanan sistem. Bahkan, terdapat indikasi pemanfaatan teknologi canggih seperti kecerdasan buatan (*Artificial Intelligence*/AI) untuk memperluas dan mempercepat serangan.

Selama ini, metode deteksi *phishing* masih banyak yang dilakukan secara tidak otomatis, seperti verifikasi manual oleh tim keamanan, pengecekan melalui daftar hitam (*blacklist*), atau pemeriksaan berdasarkan laporan pengguna. Metode tersebut memiliki keterbatasan karena memerlukan waktu yang relatif lama, bergantung pada

intervensi manusia, serta sulit menangani volume situs web yang terus bertambah. Sementara itu, deteksi otomatis memanfaatkan algoritma pembelajaran mesin yang mampu menganalisis pola data secara cepat dan mandiri. Sistem otomatis dapat mengklasifikasikan situs *phishing* dalam hitungan milidetik tanpa memerlukan pemeriksaan manual, sehingga lebih efektif digunakan dalam skala besar dan secara *real-time*.

Beberapa penelitian terdahulu telah mencoba berbagai pendekatan otomatis, mulai dari *stacking*, *Random Forest*, *Support Vector Machine* (SVM), *Iterative Dichotomiser 3* (ID3), hingga *Naïve Bayes*. Namun, penelitian-penelitian tersebut masih memiliki beberapa keterbatasan, antara lain penggunaan dataset tunggal tanpa pengujian pada variasi data lain, fokus evaluasi yang hanya mengandalkan akurasi, minimnya penanganan *class imbalance*, ketiadaan *hyperparameter tuning* yang optimal, serta jarangnya dilakukan analisis kesalahan (*error analysis*).

Selain itu, algoritma *gradient boosting* modern seperti *XGBoost* yang terbukti unggul dalam menangani data berukuran besar, tidak seimbang, dan kompleks masih jarang dimanfaatkan secara optimal untuk deteksi *phishing* otomatis. *XGBoost* (*Extreme Gradient Boosting*) dipilih dalam penelitian ini karena memiliki sejumlah keunggulan, antara lain kemampuan *regularization* untuk membantu mencegah *overfitting*, dukungan *parallel processing* untuk mempercepat proses komputasi, serta kemampuan menangani distribusi data yang tidak seimbang melalui parameter *scale_pos_weight*. Studi terbaru oleh Bahaghighat & Ghasemi (2023) menunjukkan bahwa *XGBoost* mencapai akurasi sebesar 99,2%, presisi sebesar 99,1%, dan *recall* sebesar 99,4% pada dataset berukuran besar yang terdiri atas lebih dari 88.000 situs web, dengan waktu prediksi sebesar 1500 ms atau 869 ms setelah dilakukan reduksi dimensi menggunakan *Principal Component Analysis* (PCA). Hal tersebut menunjukkan bahwa *XGBoost* memiliki performa prediktif yang tinggi sekaligus efisien untuk mendukung deteksi *phishing* dalam skala besar dan secara *real-time*.

Berdasarkan pertimbangan tersebut, penelitian ini berfokus pada pengembangan dan penerapan model *XGBoost* untuk mendeteksi situs web *phishing* secara otomatis menggunakan dataset terklasifikasi dari sumber terpercaya. Penelitian ini bertujuan

untuk menguji performa model secara komprehensif menggunakan metrik akurasi, presisi, *recall*, dan *F1-score*, serta mengidentifikasi fitur yang paling berpengaruh dalam proses klasifikasi. Hasil penelitian diharapkan dapat memberikan kontribusi nyata bagi keamanan siber, baik sebagai solusi praktis berupa prototipe deteksi *phishing* otomatis maupun sebagai referensi akademik untuk pengembangan metode serupa pada masa mendatang.

1.2 Perumusan Masalah

Berdasarkan latar belakang di atas, rumusan masalah dalam penelitian ini adalah:

1. Bagaimana cara menerapkan model XGBoost untuk mendeteksi situs web *phishing*?
2. Bagaimana performa model XGBoost dalam mengklasifikasikan situs web phishing dan situs web yang sah, diukur menggunakan metrik akurasi, presisi, *recall*, dan *F1-score*?
3. Fitur-fitur teknis apa saja yang memiliki bobot atau pengaruh terbesar dalam klasifikasi situs web phishing menggundi model XGBoost?

1.3 Batasan Masalah

Agar ruang lingkup penelitian tidak terlalu luas, batasan masalah yang diterapkan adalah sebagai berikut:

1. Model yang digunakan dalam penelitian ini adalah XGBoost (*Extreme Gradient Boosting*) sebagai metode utama dalam klasifikasi situs web *phising*.
2. Data yang digunakan dalam pelatihan dan pengujian model berasal dari dataset publik yang telah diklasifikasikan sebagai situs web *phising* atau bukan.
3. Penelitian ini hanya berfokus pada deteksi berbasis karakteristik teknis situs web yang diekstraksi dari URL dan kode sumber halaman, seperti panjang URL (*URLLength*), jumlah subdomain (*NoOfSubDomain*), rasio digit dalam URL (*DegitRatioInURL*), dan keberadaan iframe (*NoOfiFrame*). Penelitian ini tidak mempertimbangkan faktor sosial atau psikologis yang mungkin terlibat dalam serangan phishing.
4. Model yang dikembangkan tidak mencakup metode deteksi berbasis analisis teks, gambar, atau konten situs web secara *real-time*.

5. Implementasi sistem deteksi ini dilakukan dalam bentuk prototipe dan tidak langsung diintegrasikan dengan sistem keamanan *real-time* pada perangkat pengguna.

1.4 Tujuan dan Manfaat

1.4.1 Tujuan

Tujuan utama dari pengembangan sistem ini adalah untuk menciptakan solusi otomatis dalam mendeteksi situs web *phishing* melalui implementasi model *machine learning* berbasis algoritma XGBoost. Melalui sistem ini, performa model akan dianalisis dan dievaluasi secara mendalam menggunakan berbagai metrik pengujian, seperti akurasi, presisi, *recall*, dan *F1-score*, guna memastikan tingkat ketepatan yang tinggi dalam membedakan antara situs yang sah dan situs yang berbahaya. Selain itu, penelitian ini bertujuan untuk mengidentifikasi berbagai fitur teknis yang memiliki pengaruh paling signifikan dalam proses klasifikasi. Dengan demikian, sistem ini diharapkan tidak hanya mampu mendeteksi ancaman secara efektif, tetapi juga memberikan pemahaman yang lebih komprehensif mengenai karakteristik utama dari situs web *phishing*.

1.4.2 Manfaat

Hasil dari penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Meningkatkan perlindungan bagi masyarakat terhadap ancaman serangan *phishing* melalui penyediaan prototipe sistem deteksi otomatis yang akurat dan responsif.
2. Meminimalisir risiko pencurian data pribadi dengan memfasilitasi pengguna untuk melakukan identifikasi situs web berbahaya secara mandiri dan cepat.
3. Memberikan kontribusi teknis dalam pengembangan model klasifikasi berbasis XGBoost yang dievaluasi secara terukur melalui metrik akurasi dan kecepatan waktu deteksi URL.
4. Meningkatkan kesadaran keamanan digital pengguna melalui pemahaman terhadap fitur-fitur teknis yang menjadi indikator utama sebuah situs web *phishing*.
5. Menyediakan data evaluasi yang komprehensif berdasarkan pengalaman pengguna langsung terkait kemudahan penggunaan sistem dan peningkatan rasa aman saat beraktivitas di ruang siber.