BAB 1
PENDAHULUAN

1.1 Latar Belakang Masalah

Perkembangan era digital saat ini memberikan dampak yang cukup masif. Salah satu dampak atau perubahan dari perkembangan Teknologi Informasi dan Komunikasi (TIK) ini adalah situs web [1]. Situs web merupakan kumpulan halaman yang dirancang untuk menyajikan berbagai jenis informasi, seperti teks, gambar, suara, animasi, video, maupun kombinasi dari semuanya, baik yang bersifat statis maupun dinamis [2]. Kemampuannya dalam menyampaikan informasi secara cepat dan selalu diperbarui menjadikan situs web sebagai sarana yang efektif bagi setiap orang atau organisasi di berbagai wilayah untuk memperoleh informasi hanya melalui akses internet [3].

Seiring dengan semakin luasnya penggunaan situs web, platform ini juga kerap menjadi sasaran utama kejahatan siber, di mana celah keamanan yang ada sering dimanfaatkan oleh pelaku untuk melancarkan serangan siber [4]. Peningkatan penggunaan internet yang semakin besar juga memunculkan berbagai anomali dalam aktivitas digital. Arus informasi yang sangat cepat dan volume data yang terus bertambah membuka peluang bagi pelaku kejahatan siber untuk mengeksploitasi celah keamanan secara lebih efektif dan sulit terdeteksi. Anomali tersebut dapat berupa lonjakan trafik yang tidak wajar, aktivitas *login* yang mencurigakan, penyebaran tautan berbahaya secara masif, hingga munculnya pola komunikasi yang menyerupai aktivitas normal namun sebenarnya bersifat manipulatif [5]. Dalam beberapa tahun terakhir, serangan *phishing* menjadi salah satu bentuk ancaman siber yang paling menonjol dan banyak dihadapi oleh pengguna internet, baik pada sektor pemerintah, organisasi penyedia layanan, maupun masyarakat umum.

Phishing adalah ancaman keamanan yang berbahaya yang memanfaatkan teknik psikologis dan rekayasa sosial yang canggih untuk menipu individu agar mengklik tautan situs web berbahaya dan mengirimkan informasi sensitif yang sangat berharga, seperti informasi pribadi atau korporat serta kredensial akun [6]. Situs web *phishing* sering digunakan sebagai titik masuk dalam serangan rekayasa sosial secara daring dan menjadi salah satu metode yang umum dalam berbagai bentuk penipuan berbasis web [7]. Serangan *phishing* tidak terlalu kompleks secara

teknologi dan pelaksanaannya tidak memerlukan banyak usaha. Namun demikian, serangan ini umumnya sangat efektif. Penyerang membuat situs *phishing* yang dirancang dengan baik dan menyerupai tampilan serta nuansa situs resmi yang mereka tiru, sehingga sangat sulit bagi individu untuk mengenali situs *phishing* tersebut [8].

Pelaku *phishing*, yang dikenal sebagai *phisher*, umumnya menyamar sebagai pihak terpercaya seperti bank atau layanan digital untuk meyakinkan korban agar mengakses tautan berbahaya. Salah satu teknik yang paling sering digunakan adalah pembuatan situs web tiruan yang menyerupai situs asli untuk mengelabui pengguna. Melalui tautan *Uniform Resource Locator (URL)* yang dikirimkan kepada korban, pelaku berupaya mengumpulkan data pribadi secara ilegal [9]. Situs web *phishing* tersebut dibuat untuk meniru identitas organisasi resmi sehingga pengguna tidak menyadari bahwa mereka sedang mengakses halaman palsu [10]. *Phishing* bukan hanya sekadar manipulasi visual, tetapi juga telah berevolusi menjadi teknik yang sangat canggih dan sulit dideteksi. Saat ini, sekitar 80% situs web *phishing* telah menggunakan protokol *HTTPS* guna memberikan kesan aman dan sah kepada calon korban, yang secara efektif melemahkan mekanisme deteksi berbasis daftar hitam (*blacklist*) tradisional [11]. Lebih jauh lagi, integrasi kecerdasan buatan generatif oleh aktor ancaman telah memicu munculnya kampanye *phishing polymorphic*, di mana struktur *URL* dan konten situs diubah secara dinamis untuk menghindari pemindaian filter keamanan [12]. Dalam menghadapi dinamika ini, deteksi berbasis karakteristik *URL* menjadi sangat krusial. Analisis *URL* memungkinkan identifikasi ancaman secara instan di sisi klien (*client-side*) tanpa memerlukan pengunduhan konten halaman web yang berat dan berisiko bagi perangkat. Karakteristik *URL* seperti fitur leksikal (panjang *URL*, penggunaan simbol @, jumlah titik), abnormalitas struktur, serta metadata domain (usia pendaftaran dan reputasi) tetap menjadi prediktor yang sangat kuat dalam klasifikasi situs berbahaya [13].

Anti-Phishing Working Group (APWG) melaporkan bahwa jumlah serangan *phishing* pada tahun 2025 mencapai sekitar 3,8 juta kasus, meningkat dibandingkan tahun 2024 yang mencatat sekitar 3,76 juta kasus. Selain itu, sektor yang paling sering menjadi target serangan *phishing* pada kuartal keempat tahun 2025 adalah media sosial serta layanan Software-as-a-Service (SaaS) dan webmail, dengan masing-masing menyumbang 20,3% dari total serangan *phishing* yang terdeteksi. APWG juga menemukan bahwa jumlah serangan Business Email Compromise (BEC) yang melibatkan *wire transfer* meningkat drastis sebesar 136% dibandingkan kuartal sebelumnya. Pada periode yang sama, metode *phishing* melalui SMS

mengalami peningkatan, sedangkan penggunaan kode QR sebagai media *phishing* mengalami penurunan sebesar 9%. [14]. Dalam sebuah studi terkait pengalaman pengguna terhadap serangan *phishing*, ditemukan bahwa pengguna komputer dapat menjadi korban *phishing* karena lima alasan utama. Pertama, pengguna tidak memiliki pemahaman yang mendalam mengenai struktur dan karakteristik *URL*. Kedua, pengguna tidak mengetahui dengan pasti halaman web mana yang dapat dipercaya. Ketiga, pengguna sering kali tidak melihat keseluruhan alamat situs web akibat adanya pengalihan (*redirection*) atau *URL* yang disamarkan. Keempat, keterbatasan waktu membuat pengguna tidak melakukan pemeriksaan *URL* secara cermat atau secara tidak sengaja mengakses halaman tertentu. Kelima, pengguna kesulitan membedakan antara situs web *phishing* dan situs web yang sah karena tampilannya yang sangat mirip [15].

Dampak yang ditimbulkan oleh serangan *phishing* tidak hanya merugikan individu, tetapi juga organisasi apabila terjadi dalam skala besar. Kebocoran data akibat *phishing* dapat menyebabkan kerugian finansial yang signifikan serta menurunkan reputasi institusi [16]. Hal ini juga dapat merusak hak kekayaan intelektual individu maupun organisasi, serta mengancam aset keuangan pribadi dan publik. Serangan *phishing* yang berhasil dapat memicu serangan lanjutan seperti *malware*, *ransomware*, atau serangan jaringan lainnya. Berbagai jenis kejahatan siber, mulai dari serangan psikologis yang tidak berwujud hingga kerugian finansial yang nyata, dapat disebabkan oleh situs web penipuan [17].

Machine learning merupakan bagian dari *Artificial Intelligence (AI)* yang dapat "mempelajari" pola dari data pelatihan dan selanjutnya membuat kesimpulan yang akurat terhadap data baru. Kemampuan dalam mempelajari pola ini memungkinkan model *maching learning* untuk membuat keputusan atau prediksi tanpa instruksi eksplisit yang telah dikodekan secara pasti [18]. Implementasi *machine learning* untuk menyelesaikan suatu permasalahan memiliki beberapa prasyarat, yaitu membutuhkan memori yang besar, memerlukan proses pelabelan data, serta tidak jarang menghasilkan keluaran yang tidak atau kurang akurat. *Machine learning* memungkinkan komputer mempelajari pola dari sebuah data historis yang kemudian dapat digunakan untuk mengidentifikasi *website phishing* [19][20].

Algoritma *machine learning* yang populer dan terbukti efektif adalah *Random Forest* dan *Decision Tree*. Algoritma *Random* Forest bekerja dengan membangun sekumpulan pohon keputusan (*ensemble of decision trees*) yang mampu menangkap interaksi non-linear antar fitur *URL* tanpa terjebak dalam masalah *overfitting* [21].

Beberapa studi melaporkan tingkat akurasi algoritma ini mencapai lebih dari 96%, yang menunjukkan reliabilitasnya dalam skenario data yang besar dan kompleks [22][23]. Sementara itu, *Decision Tree* merupakan algoritma yang bersifat fleksibel, mudah diinterpretasikan, serta mampu mengolah hubungan non-linear dengan lebih sederhana [24]. Ide dasar di balik algoritma *Decision Tree* adalah membagi data secara rekursif menjadi beberapa subset berdasarkan nilai atribut yang berbeda hingga kriteria penghentian terpenuhi. Proses ini menghasilkan struktur pohon, di mana setiap node mewakili keputusan atau pembagian berdasarkan atribut tertentu. Penelitian mengenai deteksi *phishing* berbasis *machine learning* telah banyak dilakukan menggunakan berbagai algoritma klasifikasi, seperti *Decision Tree*, *Random Forest*, dan *Support Vector Machine* [25]. Evaluasi model umumnya dilakukan menggunakan metrik *accuracy*, *precision*, *recall*, *F1-score*, dan *confusion matrix*. Namun, sebagian penelitian masih berfokus pada nilai performa secara umum dan belum membahas secara mendalam kesalahan klasifikasi, khususnya *false negative*, yaitu kondisi ketika URL *phishing* diklasifikasikan sebagai URL *legitimate*.

Beberapa penelitian terdahulu menunjukkan bahwa *Random Forest* cenderung memberikan performa lebih baik dibandingkan *Decision Tree* karena menggunakan sejumlah pohon keputusan untuk menghasilkan prediksi yang lebih stabil. Meskipun demikian, beberapa penelitian masih menggunakan dataset berukuran terbatas dan melakukan seleksi fitur secara terpisah sebelum proses validasi silang. Kondisi tersebut dapat menyebabkan hasil evaluasi kurang mencerminkan kemampuan generalisasi model pada data baru.

Berdasarkan (*gap*) tersebut, penelitian ini membandingkan algoritma *Decision Tree* (*baseline*) dan *Random Forest* sebagai model utama. Proses seleksi fitur menggunakan *SelectFromModel* diintegrasikan bersama *SimpleImputer* dan *classifier* dalam satu *pipeline*. Selain membandingkan performa kedua model, penelitian ini juga menganalisis *feature importance* dan kesalahan *false negative* untuk mengetahui fitur yang paling berpengaruh serta risiko URL *phishing* yang lolos dari proses deteksi.

1.2 Rumusan Masalah

Serangan *phishing* melalui URL masih menjadi ancaman karena teknik yang digunakan terus berkembang dan semakin sulit dibedakan dari URL yang sah. Banyak sistem deteksi yang ada saat ini belum mampu mengidentifikasi URL

phishing secara akurat, khususnya dalam meminimalkan kesalahan klasifikasi seperti *false negative*, di mana URL berbahaya justru diklasifikasikan sebagai aman. Kesalahan ini berpotensi menimbulkan dampak yang serius karena pengguna tetap terekspos terhadap serangan tanpa adanya peringatan.

Selain itu, pemilihan algoritma klasifikasi yang tepat menjadi tantangan tersendiri, mengingat adanya *trade-off* antara interpretabilitas model dan performa prediksi. Model yang sederhana seperti *Decision Tree* mudah dipahami namun rentan terhadap *overfitting*, sementara metode *ensemble* seperti *Random Forest* menawarkan peningkatan akurasi tetapi dengan kompleksitas yang lebih tinggi. Oleh karena itu, diperlukan analisis yang lebih mendalam untuk mengetahui sejauh mana kedua algoritma ini mampu mengatasi permasalahan deteksi URL *phishing*, khususnya dalam mengurangi kesalahan klasifikasi yang kritis.

Berdasarkan permasalahan tersebut, maka dirumuskan beberapa rumusan masalah sebagai berikut:

a. Bagaimana kemampuan algoritma *Decision Tree* dalam mengklasifikasikan situs web *phishing* dan *legitimate* berdasarkan karakteristik URL, *metadata domain*, dan fitur struktur halaman web yang tersedia pada dataset PhiUSIIL sebagai model *baseline*?

b. Bagaimana perbandingan performa algoritma *Decision Tree* dan *Random Forest* berdasarkan metrik-metrik *Accuracy, Precision, Recall*, dan *F1-Score* untuk membuktikan efektivitas model?

c. Fitur karakteristik *URL* apa saja yang paling berpengaruh dalam proses klasifikasi menurut analisis *feature importance*, serta bagaimana karakteristik kasus *false negative* yang dihasilkan oleh model?

1.3 Batasan Permasalahan

Agar penelitian lebih terarah dan terfokus, batasan masalah dalam penelitian ini adalah sebagai berikut:

a. Penelitian berfokus pada analisis karakteristik URL berbasis fitur struktural domain yang tersedia pada dataset PhiUSIIL. Penelitian tidak mencakup analisis konten halaman web menggunakan *Natural Language Processing* (NLP), analisis visual seperti *screenshot* atau logo situs, maupun pengambilan halaman web secara *real-time*.

b. Proses validasi model dilakukan menggunakan metode *10-fold Stratified Cross-Validation* yang diintegrasikan dalam optimasi hyperparameter melalui *GridSearchCV*. Penelitian ini tidak membandingkan metode validasi lain seperti *bootstrapping* atau *leave-one-out cross-validation*.

c. Feature selection dilakukan menggunakan metode *Embedded Feature Selection* berbasis *Random Forest Feature Importance* dengan bantuan *SelectFromModel*. Penelitian ini tidak membandingkan metode feature selection lain seperti *filter method* maupun *wrapper method*.

1.4 Tujuan Penelitian

Tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut:

a. Mengetahui kemampuan algoritma *Decision Tree* dalam mengklasifikasikan situs web *phishing* dan *legitimate* berdasarkan karakteristik URL, *metadata domain*, dan fitur struktur halaman *web* yang tersedia pada dataset PhiUSIIL

b. Menganalisis perbandingan performa algoritma *Decision Tree* dan *RF* berdasarkan metrik *Accuracy*, *Precision*, *Recall* (khusus kelas *phishing*), dan *F1-Score* untuk membuktikan efektivitas metode *ensemble*.

c. Mengidentifikasi fitur karakteristik *URL* yang paling berpengaruh dalam proses klasifikasi serta menganalisis penyebab terjadinya *false negative* pada model deteksi *phishing*.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat secara keilmuan dan praktis sebagai berikut.

a. **Manfaat Keilmuan**

Penelitian ini diharapkan dapat menambah wawasan dan referensi dalam bidang *machine learning* dan keamanan siber, khususnya terkait penerapan algoritma *Decision Tree* dan *Random Forest* untuk mendeteksi URL *phishing*. Selain itu, hasil penelitian ini dapat memberikan gambaran mengenai perbandingan performa kedua algoritma, kontribusi fitur dalam proses klasifikasi, serta karakteristik kasus *false negative*. Hasil tersebut dapat

digunakan sebagai dasar atau referensi bagi penelitian selanjutnya yang membahas deteksi *phishing* berbasis karakteristik URL.

b. **Manfaat Praktis**

Secara praktis, penelitian ini diharapkan dapat memberikan gambaran bagi pengembang sistem, organisasi, dan pihak terkait mengenai penggunaan algoritma *machine learning* dalam mendeteksi URL *phishing*. Hasil penelitian dapat menjadi pertimbangan dalam pemilihan model klasifikasi yang memiliki performa lebih baik dan jumlah *false negative* yang lebih rendah. Selain itu, *prototype* yang dikembangkan dapat digunakan sebagai demonstrasi penerapan model klasifikasi terhadap data uji, sehingga dapat mendukung pengembangan sistem deteksi *phishing* yang lebih lanjut.

1.6 Sistematika Penulisan

a. **Bab I PENDAHULUAN**

Pada bab I berisi latar belakang penelitian, rumusan masalah, tujuan penelitian, batasan masalah, manfaat penelitian, serta sistematika penulisan laporan yang menjadi awal dalam penelitian ini. Bab ini memberikan permasalahan mengenai *link phishing* serta pentingnya penerapan metode deteksi menggunakan pendekatan *machine learning*.

b. **Bab II LANDASAN TEORI**

Pada bab II berisi landasan teori dan tinjauan pustaka yang mendukung penelitian. Pembahasan mencakup konsep dasar *cyber security*, karakteristik *link phishing*, konsep machine *learning*, algoritma *Decision Tree*, *Random Forest, feature importance, confusion matrix,* dan matriks evaluasi, serta perbandingan dengan penelitian-penelitian terdahulu yang masih relevan dengan penelitian ini sebagai referensi dan dasar pengembangan metode yang akan digunakan.

c. **Bab III METODOLOGI PENELITIAN**

Bab ini menjelaskan metodologi penelitian yang digunakan meliputi tahapan pengumpulan data, proses *preprocessing* data, ekstraksi dan seleksi fitur, pembagian data latih dan data uji, perancangan model deteksi *phishing* menggunakan algoritma *Decision Tree* dan *Random Forest*, serta evaluasi

performa model menggunakan metrik seperti *accuracy, precision, recall*, *dan F1-score.*

d. **Bab IV HASIL DAN PEMBAHASAN**

Bab ini menyajikan hasil implementasi model serta pembahasan dari penelitian yang dilakukan. Pembahasan mencakup hasil pelatihan dan pengujian model *Decision Tree* dan *Random Forest*, analisis performa model berdasarkan confusion *matrix* dan metrik evaluasi, perbandingan kinerja kedua algoritma, serta analisis kesalahan klasifikasi seperti *false negative* dalam proses deteksi *phishing*.

e. **Bab V KESIMPULAN DAN SARAN**

Bab ini berisi kesimpulan yang diperoleh dari hasil penelitian dan pembahasan yang telah dilakukan. Selain itu, bab ini juga memuat saran untuk pengembangan penelitian selanjutnya, khususnya dalam meningkatkan akurasi sistem deteksi *phishing* serta pengembangan fitur tambahan di luar karakteristik *URL* dengan memanfaatkan metode atau algoritma *machine learning* lainnya.

UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA