BAB I
PENDAHULUAN

1.1 Latar Belakang

Dengan meningkatnya penggunaan internet, situs web yang mendukung dan memfasilitasi berbagai aktivitas seperti e-commerce, perbankan, perdagangan, pembayaran tagihan, pembaruan berita, dan hiburan semakin popular. Namun, masalah muncul ketika situs web resmi dengan sumber informasi asli bersertifikat disalin dan diclone menggunakan alat dan perangkat lunak pengembangan lanjutan sehingga terlihat sangat mirip dengan situs web resmi, padahal sebenarnya bukan. Situs web semacam ini digunakan untuk tujuan menipu dan mencuri uang dari orang-orang yang tidak menyadari bahwa mereka berada di situs phishing palsu. Akibatnya, data sensitif seperti detail bank, nomor pin, nomor telepon, dan alamat bisa jatuh ke tangan yang salah [1].

Phishing adalah suatu tindakan penipuan yang menggunakan metode umpan untuk mencuri informasi rahasia dan sensitif. Serangan ini bertujuan untuk mengambil data penting seperti kata sandi, nomor, dan PIN, lalu digunakan untuk tujuan jahat. Untuk mengatasi masalah ini, algoritme pembelajaran mesin digunakan untuk mengidentifikasi URL situs web yang mencurigakan. Dengan pembelajaran mesin, kita dapat memahami berbagai jenis URL yang ada dalam daftar hitam situs web phishing, serta fitur-fitur lain yang membedakan situs web asli dan palsu. Hal ini memungkinkan algoritme pembelajaran mesin untuk memberikan tingkat akurasi yang tinggi dalam menentukan keabsahan situs web berdasarkan URL-nya. Berdasarkan laporan statistik dari Indonesia Anti-Phishing Data Exchange (IDADX) pada kuartal kedua tahun 2023, terlihat bahwa jumlah laporan serangan phishing mengalami peningkatan yang cukup besar dalam lima tahun terakhir. Pada tahun 2023, jumlah laporan serangan phishing meningkat hingga empat kali lipat lebih tinggi dibandingkan lima tahun sebelumnya. Hal ini menunjukkan adanya tren yang mengkhawatirkan terkait aktivitas phishing di Indonesia, yang telah meningkat pesat selama lima tahun terakhir. Situasi ini menegaskan pentingnya upaya untuk meningkatkan kesadaran dan perlindungan

masyarakat terhadap ancaman phishing, guna melindungi informasi pribadi dan sensitif dari serangan yang berbahaya tersebut [2].

Penelitian tentang cara mendeteksi situs web phishing untuk mengurangi kerugian dan kejahatan siber. Machine learning, yang menggunakan algoritma dan teknik khusus untuk melakukan deteksi situs web phishing. Penelitian sebelumnya telah menggunakan metode feature seleksi, di mana 20 fitur dipilih dari total 48 fitur yang ada pada situs web untuk analisis. Algoritma random forest kemudian digunakan untuk mengklasifikasikan situs web. Hasil penelitian menunjukkan bahwa menggunakan 20 fitur dan algoritma random forest, pendekatan ini berhasil menunjukkan akurasi sebesar 98,11% [3].

Oleh karena itu, dengan menggunakkan model hybrid feature seleksi seperti seperti Pearson Correlation, Information Gain, Gain Ratio, Chi-Square, Recursive Feature Elimination (RFE) dan menbahkan meta algoritma menjadi metode yang diuji kinerjanya terhadap proses deteksi web-phising.

1.2 Rumusan Masalah

Dari permasalahan yang telah dijelaskan pada bagian latar belakang, diperoleh rumusan masalah terkait bagaimana penerapan meta algoritma dalam meningkatkan kinerja model machine learning dalam deteksi web phising setelah melewati proses fitur seleksi?

1.3 Batasan Masalah

Untuk mempersempit pembahasan pada skripsi ini, maka di buat Batasan-batasan sebagai berikut :

a. Dataset menggunakan Phishing Dataset dari UCI Machine Learning Repository [44]
b. Menggunakan metode fitur seleksi yang telah diteliti pada penelitian [45] yakni Pearson Correlation, Information Gain, Gain Ratio, Chi-Square dan Recursive Feature Elimination
c. Meta algoritma menggunakan Bagging, Boosting, dan Stacking
d. Algoritma menggunakan Decision Tree untuk mengklasifikasi web-phising

e. Metrik evaluasi menggunakan akurasi untuk menghitung kinerja dari model

1.4 Tujuan Penelitian

Tujuan yang ingin dicapai dalam pembuatan laporan skripsi ini adalah "Menerapkan teknik meta algoritma untuk mengetahui adanya peningkatan kinerja klasifikasi dalam proses deteksi web phising setelah melewati proses fitur seleksi"

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat untuk peneliti tentang proses deteksi web phishing, diantaranya :

a. Penelitian ini diharapkan bisa bermanfaat untuk pengembangan deteksi web-phising di masa depan
b. Bagi peneliti, penelitian ini digunakan untuk mengetahui pengaruh model hybrid dalam proses deteksi web-phising, sehingga menghasilkan kinerja yang utuh