BAB I

PENDAHULUAN

1.1 Latar Belakang

Evolusi teknologi digital modern yang berkelanjutan telah mendorong ketergantungan masif terhadap penggunaan aplikasi web dalam bidang finansial, medis, edukasi, dan birokrasi. Aplikasi web kini menjadi tulang punggung operasional bagi banyak organisasi karena menyediakan layanan yang mudah diakses dan efisien bagi pengguna. Namun, peningkatan adopsi teknologi ini juga diiringi dengan peningkatan serangan siber yang semakin variatif, padat dan sulit dikendalikan. Berdasarkan laporan [1], rata-rata biaya kebocoran data global mencapai USD 4,88 juta per insiden, menandakan bahwa serangan siber telah menjadi ancaman serius terhadap keberlangsungan bisnis dan pelayanan publik. Lonjakan serangan siber sebesar 32% pada tahun 2023, dengan total 361 juta anomali menurut laporan BSSN, mempertegas fakta bahwa pertahanan digital Indonesia saat ini menghadapi risiko eksploitasi yang semakin serius [2].

Serangan terhadap aplikasi web menjadi salah satu bentuk ancaman yang umum untuk terjadi di berbagai *platform* digital. Jenis serangan siber yang paling marak ditemukan menurut [3] meliputi SQL *Injection*, XSS, *Broken Authentication*, dan *Remote Code Execution*. Ancaman-ancaman tersebut sering kali mengeksploitasi kelemahan pada sistem validasi *input* atau otentikasi pengguna. Dalam konteks ini, mekanisme pertahanan seperti *Web Application Firewall* (WAF) menjadi lapisan keamanan penting yang berfungsi untuk menyaring dan memantau permintaan HTTP guna mencegah aktivitas berbahaya. Namun, pendekatan

tradisional berbasis tanda tangan (*signature-based detection*) atau berbasis aturan (*rule-based detection*) yang umum digunakan dalam WAF konvensional memiliki keterbatasan dalam mendeteksi *zero-day attack* atau variasi serangan baru yang belum memiliki pola tanda tangan atau aturan tertentu.

Keterbatasan WAF konvensional tersebut dapat diatasi dengan pendekatan baru yaitu WAF berbasis *machine learning* yang mampu mempelajari pola perilaku lalu lintas HTTP normal dan mengenali anomali yang berpotensi sebagai serangan. Selain menggunakan pola tanda tangan, pendekatan ini juga menggunakan kemampuan model dalam beradaptasi dengan pola serangan baru. Sejumlah penelitian terdahulu telah menunjukkan efektivitas pendekatan ini. Penelitian oleh [4] mengembangkan WAF berbasis *anomaly detection* menggunakan teknik *n-gram* dan TF-IDF pada *dataset* CSIC 2010, dan memperoleh *accuracy* hingga 99,53%, yang membuktikan efektivitas tinggi *machine learning* dalam mendeteksi serangan seperti SQL *Injection* dan XSS. Selanjutnya, penelitian oleh [5] menggunakan pendekatan *feature engineering* dengan algoritma *Decision Tree* dan SVM, menghasilkan *accuracy* 99,6% dengan *false positive rate* serendah 0,54%, menunjukkan potensi integrasi analisis fitur dalam meningkatkan performa WAF.

Penelitian lain oleh [6] menegaskan efektivitas *Deep Learning* berbasis LSTM yang mampu mencapai *accuracy* di atas 99% dengan *F1-score* 99,4%, membuktikan kemampuan model dalam mengenali pola karakter sekuensial pada permintaan HTTP. [7] kemudian mengusulkan model *Deep Learning Web Attack Detector* (DL-WAD) berbasis Bi-LSTM dan *Natural Language Processing* (NLP) yang mencapai *accuracy* hingga 99,91%, dengan *false positive rate* rendah pada

dataset HttpParams dan CSIC 2010. Sementara itu, penelitian oleh [8] menunjukkan bahwa penerapan algoritma CNN, *Decision Tree*, dan SVM dapat mendeteksi beragam serangan siber seperti SQLi, XSS, dan *Phishing* dengan kemampuan deteksi tingkat tinggi dan efisiensi komputasi yang baik.

Kesimpulan yang dapat diambil berdasarkan berbagai studi tersebut adalah bahwa algoritma berbasis pohon keputusan (*Tree-Based Classifier*) menunjukkan performa yang konsisten dan kompetitif dalam mendeteksi serangan siber. Keunggulan algoritma ini adalah mampu dalam menangani data dalam jumlah besar, menangani risiko *overfitting* melalui mekanisme *ensemble learning*, serta menghasilkan hasil prediksi yang dapat dijelaskan secara interpretatif (*explainable model*).

Penelitian ini berfokus pada pengembangan dan perbandingan keempat algoritma *Tree-Based Classifier* tersebut untuk membangun model *Web Application Firewall* berbasis *machine learning* menggunakan *dataset* CSIC 2010. Penelitian ini bertujuan untuk membangun dan menghasilkan model WAF berbasis *machine learning* yang paling efektif dan efisien, serta mampu mendeteksi serangan web secara otomatis dengan tingkat *false positive* yang rendah, sehingga dapat berkontribusi terhadap penguatan sistem keamanan aplikasi web di Indonesia.

1.2 Rumusan Masalah

Berdasarkan latar belakang tersebut, maka masalah yang akan diteliti dalam penelitian ini adalah:

1. Bagaimana membangun dan menghasilkan model *Web Application Firewall* (WAF) berbasis *machine learning* yang mampu

mengklasifikasikan lalu lintas HTTP ke dalam kategori normal dan anomali (serangan) dengan hasil metrik evaluasi yang tinggi menggunakan *dataset* CSIC 2010?

2. Bagaimana perbandingan kinerja empat algoritma *Tree-Based Classifier* dalam melakukan klasifikasi biner serangan siber pada lalu lintas HTTP dengan menggunakan metrik evaluasi?

1.3 Tujuan Penelitian

Bertitik tolak dari permasalahan yang telah dipaparkan, sasaran utama yang ingin dicapai dalam penelitian ini adalah sebagai berikut:

1. Mengembangkan model *Web Application Firewall* (WAF) berbasis *machine learning* yang mampu mengklasifikasikan lalu lintas HTTP ke dalam kategori normal dan anomali (serangan) secara otomatis menggunakan *dataset* CSIC 2010 dengan hasil metrik evaluasi yang tinggi.
2. Menganalisis dan membandingkan kinerja empat algoritma *Tree-Based Classifier* dalam klasifikasi biner serangan siber pada lalu lintas HTTP dengan menggunakan metrik evaluasi untuk menentukan algoritma yang paling optimal.

1.4 Batasan Masalah

Untuk memfokuskan penelitian, masalah dalam penelitian ini dibatasi sebagai berikut:

1. Penelitian ini menggunakan *dataset* CSIC 2010 HTTP *Dataset*, yang terdiri dari 36.000 permintaan HTTP normal dan 25.065 permintaan

HTTP anomali. *Dataset* ini mencakup berbagai serangan pada *web application*, yang dihasilkan dari simulasi aplikasi web *e-commerce*.

2. Penelitian ini hanya berfokus pada klasifikasi biner, yaitu mengelompokkan setiap permintaan HTTP ke dalam dua kelas, yaitu normal dan anomali (serangan). Penelitian ini tidak melakukan klasifikasi lebih lanjut terhadap jenis atau kategori serangan.
3. Ruang lingkup algoritma dalam penelitian ini dibatasi pada empat model *Tree-Based Classifier* (DT, RF, GBM, dan *XGBoost*), dengan fokus komparasi pada aspek efektivitas klasifikasi dan efisiensi komputasi masing-masing model.
4. Penilaian kinerja model dilakukan melalui analisis metrik evaluasi di antaranya *accuracy*, *precision*, *recall*, dan *f1-score* berdasarkan *confusion matrix*. Perlu dicatat bahwa pengujian ini terbatas pada penggunaan data uji dari *dataset* CSIC 2010 dan tidak mencakup implementasi pada lalu lintas HTTP secara nyata dalam lingkungan produksi.

1.5 Manfaat Penelitian

Model *Web Application Firewall* (WAF) berbasis *Machine Learning* yang dikembangkan dapat memberikan manfaat dengan harapan sebagai berikut:

1. Bagi Pemilik Usaha atau Layanan Digital:
    a. Menyediakan solusi keamanan otomatis yang mampu mendeteksi serta memblokir berbagai bentuk serangan pada *web application*, tanpa bergantung pada pembaruan tanda tangan (*signature updates*)

secara manual.

b. Mengurangi potensi kerugian finansial dan gangguan operasional akibat serangan siber, sekaligus meningkatkan keandalan sistem serta kepercayaan pengguna terhadap layanan digital yang dikelola.

2. Bagi Pengguna atau *End-User*:

a. Menjamin perlindungan terhadap data pribadi dan transaksi digital dengan menekan risiko kebocoran informasi akibat eksploitasi celah keamanan pada aplikasi web.

b. Meningkatkan pengalaman penggunaan layanan digital yang aman dan andal melalui penerapan sistem WAF cerdas yang mampu beradaptasi terhadap pola serangan baru secara dinamis dan *real-time*.

3. Bagi Ilmu Komputer dan Akademisi:

a. Penelitian ini diharapkan dapat memberikan sumbangsih teoritis dan praktis bagi pengembangan sistem deteksi ancaman siber, terutama melalui pengkajian mendalam terhadap efektivitas algoritma *Tree-Based Classifier* dalam memproteksi aplikasi web.

b. Menjadi referensi bagi penelitian selanjutnya di bidang *cybersecurity* dan *artificial intelligence*, terutama dalam pengembangan model klasifikasi adaptif yang dapat diterapkan pada sistem pertahanan siber modern.

1.6 Sistematika Penulisan

Secara garis besar, sistematika pembahasan dalam Tugas Akhir ini diuraikan

melalui bab-bab berikut:

BAB I PENDAHULUAN

Bab ini menguraikan fondasi penelitian yang mencakup latar belakang, rumusan masalah, serta tujuan dan manfaat penelitian. Selain itu, bagian ini juga memaparkan batasan masalah, sistematika penulisan, serta daftar istilah dan singkatan yang digunakan.

BAB II LANDASAN TEORI

Bab ini menyajikan tinjauan literatur yang mencakup studi-studi terdahulu yang relevan, analisis komparatif antar-penelitian, serta landasan teoritis yang menjadi dasar fundamental dalam pelaksanaan penelitian ini seperti *cyber attack*, *cybersecurity*, *firewall* dan WAF, *machine learning*, *pre-processing data*, *cleaning data*, *feature engineering*, *transformation* dan *scaling*, *feature selection* dan *feature importance*, *handling imbalance data* atau SMOTE, *cross validation*, *tree-based classifier*, *hyperparameter tuning*, evaluasi kinerja model, *jupyter notebook*.

BAB III METODOLOGI PENELITIAN

Bab ini menguraikan secara sistematis tahapan metodologi yang ditempuh dalam penelitian ini, seperti karakteristik data, *pre-processing data*, pemodelan, *hyperparameter tuning*, evaluasi kinerja model, perbandingan model, dan definisi operasional.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan pemaparan hasil eksperimen beserta analisis mendalam terhadap data yang diperoleh, merujuk pada tahapan-tahapan yang telah ditetapkan dalam metodologi penelitian.

BAB V PENGUJIAN DAN ANALISIS HASIL

Bab ini menyajikan rincian hasil pengujian serta analisis performa dari berbagai model *Tree-Based Classifier* yang telah dikembangkan, guna mengukur efektivitas dan efisiensinya dalam mendeteksi serangan.

BAB VI PENUTUP

Bab ini memaparkan poin-poin kunci sebagai kesimpulan dari seluruh rangkaian pengujian dan memberikan arahan pengembangan guna menyempurnakan batasan-batasasn yang ada dalam penelitian ini.

1.7 Daftar Istilah dan Singkatan

Daftar istilah dan singkatan dapat dilihat pada Tabel 1.1.

Tabel 1.1. Daftar Istilah dan Singkatan

Istilah/Singkatan	Keterangan
Accuracy	Metrik ini mengukur tingkan keberhasilan model dalam mengklasifikasikan seluruh sampel secara benar, baik pada kelas serangan maupun kelas normal, dibandingkan dengan total observasi.
Anomali	Lalu lintas data yang tidak normal atau mencurigakan yang berpotensi menjadi serangan siber.
BSSN	Badan Siber dan Sandi Negara.
Case Folding	Proses transformasi teks di mana semua karakter alfabet diubah menjadi huruf kecil untuk mengurangi redudansi fitur.

Confusion Matrix	Tabel evaluasi yang membandingkan hasil prediksi model dengan label asli.
CSIC	*Consejo Superior de Investigaciones Científicas* (Lembaga penelitian Spanyol pengembang *dataset* CSIC HTTP 2010).
Deduplikasi	Proses penghapusan data duplikat agar tidak terjadi redudansi.
DT	*Decision Tree* (Model prediksi berbentuk struktur pohon keputusan).
Ensemble Learning	Teknik penggabungan beberapa model *machine learning* untuk meningkatkan performa prediksi.
Entropi	Ukuran ketidakteraturan atau keacakan karakter dalam suatu *string* teks (URL atau *content*).
F1-score	Metrik evaluasi untuk mengukur keseimbangna kinerja model, rata-rata harmonik digunakan antara *precision* dan *recall*.
Feature Engineering	Proses menciptakan fitur baru atau mengubah fitur yang ada agar lebih mudah dipahami oleh model.
Feature Importance	Teknik untuk menentukan fitur mana yang paling berpengaruh terhadap hasil prediksi model.
FN (*False Negative*)	Serangan yang salah dideteksi sebagai lalu lintas normal (kecolongan).

FP (*False Positive*)	Lalu lintas normal yang salah dideteksi sebagai serangan (salah lapor).
GBM	*Gradient Boosting Machine* merupakan algoritma *enxemble* yang bekerja secara sekuensial dengan cara membangun model baru untuk meminimalkan residu atau kesalahan prediksi dari model-model sebelumnya.
HTTP	*Hypertext Transfer Protocol* (Protokol komunikasi untuk pengiriman data di web).
Hyperparameter Tuning	Proses mencari kombinasi parameter terbaik untuk mengoptimalkan kinerja model.
K-Fold Cross Validation	Metode validasi dengan membagi kumpulan data menjadi bagian *K* untuk memastikan stabilitas hasil.
Min-Max Scaling	Teknik normalisasi data untuk mengubah rentang nilai fitur menjadi 0 hingga 1.
Missing Value	Data yang hilang atau kosong atau tidak ada dalam *dataset*.
One-Hot Encoding	Proses mengubah data kategorikal menjadi representasi angka biner (0 dan 1).
Overfitting	Merupakan suatu kondisi di mana model memiliki kompleksitas yang terlalu tinggi sehingga terlalu menyesuaikan diri dengan derau pada data latih, yang mengakibatkan penurunan kemampuan generalisasi

	terhadap data baru yang belum pernah dilihat sebelumnya.
Payload	Bagian dari data transmisi yang berisi pesan atau muatan yang sebenarnya (Sering kali menjadi tempat disisipkannya serangan).
Precision	Rasio prediksi positif yang sangat akurat.
Recall	Kemampuan model dalam menemukan atau menjaring kembali seluruh data positif (serangan).
Regex	*Regular Expression* (Metode untuk mencari atau mengekstraksi pola teks tertentu).
RF	*Random Forest* (Kumpulan algoritma dengan banyak pohon keputusan).
SMOTE	Metode pengambilan sampel lebih banyak dari minoritas sintetis sebagai cara untuk menangani data yang tidak seimbang dengan menggunakan data sintetis dari kelas minoritas.
SQL *Injection* (SQLi)	Teknik serangan siber dengan memasukkan perintah SQL yang berbahaya ke dalam kolom *input* aplikasi web.
TN (*True Negative*)	Kondisi di mana data aktivitas normal berhasil diklasifikasikan secara tepat oleh model sebagai aktivitas normal.

TP (*True Positive*)	Kondisi di mana data yang berisi aktivitas serangan berhasil diklasifikasikan secara akurat oleh model ke dalam kategori serangan.
Tree-Based Classifier	Kelompok algoritma klasifikasi yang menggunakan struktur pohon keputusan dalam proses logikanya.
WAF	*Web Application Firewall* (Sistem keamanan yang menyaring dan memantau lalu lintas HTTP antara aplikasi web dan internet).
XGBoost	*Extreme Gradient Boosting* (Optimasi *gradient boosting* yang dirancang untuk kecepatan dan efisiensi tinggi).
XSS	*Cross-Site Scripting* (Serangan injeksi skrip pada sisi klien melalui halaman web).