202010370311414
Padang Ikhrosim
Prodi Informatika

BAB I

PENDAHULUAN

1.1 Latar Belakang

DDoS (*Distributed Denial of Service*) merupakan masalah lama dalam jaringan yang terus menjamur dan berkembang[1]. Serangan DDoS merupakan salah satu ancaman yang berbahaya, menurut *cloudflare* telah dilaporkan serangan terbesar yang pernah tercatat dengan kecepatan 7,3 Tbps pada Mei 2025, dan *Amazon Web Services* (AWS) melaporkan bahwa pada Q1 2020 menahan serangan DDoS sebesar 2,3 Tbit/s d [2]. DDOS merupakan upaya serangan terhadap server di dalam jaringan internet dengan cara membanjiri banyak data pada lalu lintas jaringan untuk mengganggu jalannya lalu lintas normal pada server. Serangan ini dapat mengakibatkan server menjadi *down* dan mengakibatkan *system error* [3] [4]. Serangan DDoS berdampak pada kerugian yang sangat besar, seperti berkurangnya pendapatan, kegagalan produksi, buruknya reputasi, pencurian dan lain-lain [5].

Serangan DDoS terdiri dari *Low-Rate Distributed Denial of Service* (LRDDoS) dan *High-Rate Distributed Denial of Service* (HRDDoS) [6]. serangan *High-Rate Distributed Denial of Service* (HRDDoS) dilakukan dengan mengirimkan lalu lintas jaringan dalam jumlah besar guna memanfaatkan kapasitas jaringan. Kelemahan utama dari serangan HRDDoS terletak pada karakteristik lalu lintasnya, sehingga para penyerang lebih memilih menggunakan pendekatan *Low-Rate* DDoS (LRDDoS) [7].

Serangan LRDDoS sulit dideteksi karena lalu lintas serangan menyerupai lalu lintas normal. Alih-alih menghabiskan *bandwicth* dan sumber daya jaringan, serangan LRDDoS menargetkan kelemahan pada lapisan protokol. Penyerang mengirimkan paket berbahaya dengan laju yang rendah, sehingga sistem keamanan berbasis jaringan tidak dapat mengenali ciri khas serangan tersebut [8]. Oleh karena itu, LRDDoS menjadi perhatian utama dalam masalah keamanan SDN, sebab itu

diperlukan sebuah sistem yang mampu mengidentifikasi dan mendeteksi serangan tersebut [9].

Penelitian sebelumnya, telah diusulkan beberapa metode untuk melakukan deteksi serangan LRDDoS. Pada penelitian [10] telah dibuat sistem untuk mendeteksi LRDDoS dengan metode *machine learning* menggunakan algoritma *K-Nearest Neighbors* (KNN). Jumlah dataset yang digunakan dalam penelitian ini adalah 160.006 data training dan 39.994 data test. Hasil eksperimen untuk 10 pps ini mendapatkan akurasi, presisi, recall, dan skor F1 sebesar 50.51% dan *prediction loss* sebesar 1.682%. Hasil eksperimen untuk 20 pps ini mendapatkan akurasi, presisi, recall, dan skor F1 sebesar 92.71% dan *prediction loss* sebesar 82,67%. Hasil eksperimen untuk 50 pps ini mendapatkan akurasi, presisi, recall, dan skor F1 sebesar 13.21% dan *prediction loss* sebesar 82,48%. Hasil eksperimen untuk 100 pps ini mendapatkan akurasi, presisi, recall, dan skor F1 sebesar 58.47% dan *prediction loss* sebesar 99,27%. Hasil eksperimen untuk 200 pps ini mendapatkan akurasi, presisi, recall, dan skor F1 sebesar 62.28% dan *prediction loss* sebesar 99,27%.

Pada penelitian [11] telah dibuat sistem deteksi LRDDoS dengan menggunakan algoritma *Random Forest* dengan *Logistic Regression Coefficient*. Jumlah dataset yang digunakan pada penelitian ini adalah 160.006 data training dan 39.994 data test. Hasil eksperimen untuk 50 *Packet Sending Rate* (pps) ini mendapatkan akurasi, presisi, recall, dan skor F1 sebesar 92.3% dan *prediction loss* sebesar 98,5%. Hasil eksperimen untuk 100 pps ini mendapatkan akurasi, presisi, recall, dan skor F1 sebesar 98.2% dan *prediction loss* sebesar 98,8%. Hasil eksperimen untuk 200 pps ini mendapatkan akurasi, presisi, recall, dan skor F1 sebesar 98.7% dan *prediction loss* sebesar 99.1%.

Pada penelitian [12] telah dibuat sistem deteksi LRDDoS dengan metode *machine learning*. Pada penelitian ini, Jumlah dataset yang digunakan ada 2 yaitu 420.000 data training serta 18.000 data test dan 420.000 data training serta 36.000 data test. Hasil eksperimen dengan data test 18.000 dengan algoritma SVM (RBF) mendapatkan akurasi, presisi, recall, dan skor F1 sebesar 100%. Hasil eksperimen

menggunakan algoritma SVM (LIN) mendapatkan akurasi, presisi, recall, dan skor F1 sebesar 100%. Hasil eksperimen menggunakan algoritma KNN mendapatkan akurasi, presisi, recall, dan skor F1 sebesar 87.8%, 93.1%, 88.2%, dan 86.5%. Hasil eksperimen menggunakan algoritma DTC mendapatkan akurasi, presisi, recall, dan skor F1 sebesar 100%. Hasil eksperimen menggunakan algoritma RFC mendapatkan akurasi, presisi, recall, dan skor F1 sebesar 84.2%, 86.4%, 84.6%, dan 83.1%. Hasil eksperimen menggunakan algoritma MLP mendapatkan akurasi, presisi, recall, dan skor F1 sebesar 34.5%, 32.5%, 36.3%, dan 23%. Hasil eksperimen menggunakan algoritma GNB mendapatkan akurasi, presisi, recall, dan skor F1 sebesar 93.9%, 95%, 94.1%, dan 93.5%. Sedangkan hasil eksperimen dengan data test 36.000 dengan algoritma SVM (RBF) mendapatkan akurasi, presisi, recall, dan skor F1 sebesar 100%. Hasil eksperimen menggunakan algoritma SVM (LIN) mendapatkan akurasi, presisi, recall, dan skor F1 sebesar 100%. Hasil eksperimen menggunakan algoritma KNN mendapatkan akurasi, presisi, recall, dan skor F1 sebesar 95.6%, 89.2%, 92.9%, dan 87.5%. Hasil eksperimen menggunakan algoritma DTC mendapatkan akurasi, presisi, recall, dan skor F1 sebesar 89.7.6%, 78.4.2%, 83.3%, dan 80.4%. Hasil eksperimen menggunakan algoritma RFC mendapatkan akurasi, presisi, recall, dan skor F1 sebesar 80.7%, 88.4%, 84.4%, dan 82.5%. Hasil eksperimen menggunakan algoritma MLP mendapatkan akurasi, presisi, recall, dan skor F1 sebesar 65.3%, 52.4%, 66.6%, dan 54.2%. Hasil eksperimen menggunakan algoritma GNB mendapatkan akurasi, presisi, recall, dan skor F1 sebesar 59.1%, 68.6%, 83.1%, dan 70%.

Pada penelitian [13] mengevaluasi deteksi serangan *Low-Rate Distributed Deniial of Service (LRDDoS)* pada jaringan SDN-Enabled IoT menggunakan *machine learning dan featuring impotance*. Ada tiga *feature importance* yang digunakan untuk menyederhanakan jumlah fitur antara lain : *Logistic Regression* (LR), *Random Forest Classifier* (RFC), dan *Random Forest Regresion* (RFR). Kemudian diuji menggunakan delapan algoritma yakni, SVM (LIN), SVM (RBF), RF, DT, MLP, GNB, ADB, KNN. Kemudian dilakukan uji coba dengan tiga kecepatan yaitu, 20,50,dan 70 pps. Hasil pada metode LR, hampir semua algoritma

yang diuji seperti SVM (LIN), DT, MLP, GNB, ADB dapat mencapai akurasi, presisi, *recall*, dan *F1 score* sebanyak 100% dengan *classification* sebanyak 4%. Algoritma GNB menjadi hasil yang paling unggul karena membarikan hasil klasifikasi yang paling sempurna. Sedangkan pada algoritma KNN dan SVM (RBF) menunjukkan hasil yang paling buruk dengan akurasi hanya sekitar 51.88%, presisi 25.94%, *recall* 50%, dan *F1-score* 34.16%. Pada metode RFC menunjukkan hasil yang kurang stabil terutama pada kecepatan 70 pps menunjukkan *classification loss* mencapai 51.21%. meskipun demikian algoritma seperti SVM (LIN), DT,GNB,dan ADB tetap mempertahankan performa yang sempurna dengan nilai akurasi, presisi, *recall*, dan *F1-score* mencapai 100%. GNB menujukkan performa meningkat lebih tinggi yaitu 51,88% di 20 pps, dan 100 % di 50 pps dan 70 pps. Pada metode RFR hampir semua algoritma berhasil mencapai di nilai 100% pada seluruh tingkat kecepatan paket. Algoritma GNB kembali menjadi yang tercepat dengan waktu yang dibutuhkan hanya 0.013 detik. Hal ini menunjukkan efisiensi yang luar biasa baik dalam performa maupun sumber daya. Pada metode ini, algoritma RF dan DT mencatat hasil yang sedikit lebih rendah yaitu akurasi 90.35%, recall 89.98%, dan *F1-Score* 90.17%.

Berdasarkan penelitian[14], membahas tentang pengembangan sebuah *framework Software-defined Networking* (SDN) yangn cerdas dan aman untuk lingkungan *Internet of Things* (IoT) dengan mengintegrasikan algoritma *Machine Learning*, yaitu *LightGBM* untuk optimasi *routing* dan XGBoost untuk deteksi serangan secara *real-time*. Permasalahan utama yang diangkat adalah Permasalahan utama yang diangkat adalah keterbatasan *controller* SDN konvensional dalam menangani kompleksitas trafik IoT yang dinamis serta ancaman keamanan yang terus berkembang. Oleh karena itu, penulis mengusulkan pendekatan berbasis *machine learning* yang mampu menggabungkan aspek performa jaringan dan keamanan dalam satu sistem terpadu. Metodologi penelitian menggunakan dataset sintetis sebanyak 10.000 data trafik yang mencerminkan kondisi jaringan IoT, dengan berbagai parameter seperti *latency*, *throughput*, *packet loss,* dan indikator keamanan. Hasil pengujian menunjukkan bahwa model XGBoost mampu mencapai akurasi deteksi serangan hingga 99,8% dengan *false positive rate* sebesar 0,2%,

sementara model LightGBM memberikan akurasi tinggi dalam menentukan jalur routing optimal. Selain itu, implementasi sistem ini juga terbukti meningkatkan performa jaringan secara signifikan, seperti penurunan *latency* hingga 49,7%, peningkatan *throughput* sebesar 42,7%, serta penurunan *packet loss* sebesar 75,9%. Dengan waktu respon *controller* yang hanya 23,7 ms, sistem ini dinilai mampu bekerja secara *real-time*. Secara keseluruhan, penelitian ini menunjukkan bahwa integrasi *machine learning* dalam SDN mampu menghasilkan sistem jaringan yang lebih adaptif, efisien, dan aman, meskipun masih memiliki tantangan seperti kebutuhan komputasi yang tinggi dan perlunya pembaruan model secara berkala.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk menembangkan model deteksi serangan LRDDoS dengan menggabungkan KNN dan XGBoost menggunakan metode *Ensemble Stacking* pada jaringan SD-IoT menggunakan dataset yang sama pada penelitian [13] yang didapatkan dari website Mendeley Data dengan judul LRDDoS dataset (CoAP) [15]. Diharapkan penelitian ini dapat meningkatkan akurasi pada penelitian sebelumnya menggunakan metode yang diusulkan.

1.2 Rumusan Masalah

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

1) Bagaimana cara mendeteksi serangan LRDDoS menggunakan metode *Ensemble Stacking* dengan menggabungkan KNN dan XGBoost?
2) Apakah deteksi serangan LRDDoS menggunakan metode Ensemble Stacking dengan menggabungkan KNN dan XGBoost ini bisa efektif, jika dinilai menggunakan variabel *accuracy, precision, recall, f1-score,* dan *predict loss*?

1.3 Tujuan Penelitian

Tujuan penelitian ini adalah :

1) Mengetahui hasil deteksi serangan dari *Low-Rate* DDoS
2) Mengetahui hasil *accuracy*, *precision*, recall, *f1-score* dan *prediction loss* menggunakan metode *Ensemble Stacking*.

1.4 Batasan Penelitian

Batasan masalah dari penelitian ini yaitu :

1) Dataset yang digunakan pada penelitian ini adalah LRDDoS dataset (CoAP) dari website Mendeley Data yang berjumlah 200.000 data (160.006 data train dan 39.994 data test) [15]
2) Menggunakan metode *Ensemble Stacking* dengan *K-Nearest Neighbor (KNN)* dan *eXtreme Gradient Boosting (XGBoost)*
3) Menggunakan bahasa pemrograman python.
4) Menggunakan mininet untuk emulasi jaringan di SDN
5) Menggunakan ryu controller
6) Menggunakan sistem operasi Linux
7) Pengujian model hanya dilakukan pada tiga tingkat packet rate: 20 pps, 50 pps, dan 70 pps.
8) Pengujian performa dibatasi pada parameter evaluasi: *accuracy, precision, recall, F1-score,* dan *predict loss.*
9) Penelitian ini tidak membahas mitigasi serangan, melainkan fokus pada proses deteksi.