BAB 2
LANDASAN TEORI

2.1 *Distributed Denial of Service* (DDoS)

Serangan *Distributed Denial of Service* (DDoS) telah menjadi ancaman serius dalam dunia siber yang dapat menyebabkan kerugian signifikan terhadap layanan dan operasional jaringan. Serangan DDoS melibatkan upaya yang dilakukan oleh pihak jahat untuk membuat layanan tidak tersedia bagi pengguna dengan melibatkan sejumlah besar perangkat yang membanjiri target dengan lalu lintas palsu atau memanfaatkan kelemahan tertentu[15]. Penting untuk memiliki sistem deteksi yang efektif untuk merespons dan mengurangi dampak serangan DDoS. Metode deteksi anomali menjadi pilihan utama karena dapat mengidentifikasi perilaku yang tidak biasa dalam lalu lintas jaringan.

2.2 Anomali

Anomali merupakan sesuatu yang tidak biasa atau menyimpang dari norma atau ekspektasi. Dalam konteks ilmu pengetahuan dan teknologi, anomali mengacu pada titik data atau pola yang berbeda secara signifikan dari sebagian besar data lainnya. Deteksi anomali telah menjadi bidang penelitian yang aktif di berbagai komunitas penelitian selama beberapa dekade namun masih terdapat beberapa kompleksitas masalah dan tantangan unik yang memerlukan pendekatan tingkat lanjut. Deteksi anomali penting untuk mengidentifikasi kesalahan, penipuan, atau ketidaksesuaian[16].

2.3 *Machine Learning*

Machine Learning atau pembelajaran mesin adalah bidang ilmu yang mempelajari algoritma dan model statistik yang memungkinkan sistem komputer melaksanakan tugas-tugas tertentu tanpa instruksi pemrograman secara eksplisit. *Machine learning* merupakan teknologi yang terus berkembang dan memiliki potensi besar dalam berbagai bidang seperti kesehatan, finansial, pemasaran, dan jaringan. *machine learning* digunakan untuk mengajarkan mesin cara menangani data dengan lebih efisien sehingga dapat menangani data dalam jumlah besar dan kompleks, serta mampu membuat prediksi yang akurat jika dilatih dengan baik.[17]

2.4 *Supervised vs Unsupervised Machine Learning*

Dalam pembelajaran mesin terdapat dua pendekatan utama yaitu adalah pembelajaran terawasi (*supervised learning*) di mana algoritma dilatih menggunakan dataset yang berlabel. Dataset tersebut terdiri dari pasangan *input* dan *output*,dengan catatan *output* yang benar (label) sudah diketahui sebelumnya. Algoritma bertujuan untuk mempelajari hubungan atau fungsi yang memetakan *input* ke *output*. Beberapa algoritma yang dikategorikan sebagai *supervised learning* adalah *Decision Trees*, *Neural Network*, dan *Linear Reggresion*. Metode ini biasanya digunakan untuk kepentingan klasifikasi dan regresi. Sedangkan pembelajaran tanpa pengawasan (*unsupervised learning*) merupakan metode di mana algoritma dilatih menggunakan dataset yang tidak berlabel. Algoritma berusaha menemukan struktur, pola, atau distribusi dalam data tanpa panduan output yang diketahui. Beberapa algoritma dari *unsupervised learning* adalah *Isolation Forest*, *K-Means*, dan *Principal Component Analysis* (PCA). Metode ini memiliki kegunaan yang berbeda dengan *supervised learning* seperti untuk deteksi anomali, penyederhanaan data, dan *clustering*. Kedua metode ini memiliki tujuan, teknik, dan aplikasi yang berbeda.[18]

2.5 *Isolation Forest*

Algoritma *Isolation Forest* menjadi perhatian utama dalam deteksi anomali karena kemampuannya yang baik untuk mengisolasi instansi atau titik data yang langka dalam waktu yang relatif singkat. Pendekatan ini bekerja dengan membangun pohon isolasi secara acak dan mengukur seberapa cepat titik data dapat diisolasi dalam pohon tersebut. Keunggulan algoritma ini termasuk kemampuan untuk menangani data tinggi dimensi dan efisien dalam mendeteksi anomali[19]. Algoritma *Isolation Forest* pernah diimplementasi dalam mendeteksi serangan DDoS pada lalu lintas jaringan. Hasilnya menunjukkan bahwa *Isolation Forest* mampu mengenali pola lalu lintas yang tidak biasa yang sering terjadi selama serangan DDoS. Penelitian ini memberikan dasar yang kuat untuk pengaplikasian algoritma *Isolation Forest* dalam konteks deteksi serangan DDoS[20].

2.6 *One-Class Support Vector Machine*

Support Vector Machines (SVM) dapat digunakan secara efektif untuk mendeteksi anomali, terutama saat menangani data berdimensi tinggi. *Support*

Vector Machines (SVM) berfungsi untuk mendeteksi anomali terutama melalui varian khusus yang disebut *One-Class SVM*. *One-Class SVM* adalah algoritma pembelajaran mesin tanpa pengawasan yang dilatih hanya pada kelas ”normal”. SVM akan mencoba menemukan batas yang merangkum data normal, sehingga mengidentifikasi apa pun yang berada di luar batas ini sebagai anomali.[21]. Salah satu algoritma canggih yang dapat digunakan untuk mendeteksi anomali adalah *One Class Support vector machine* (OCSVM). Efisiensi algoritma OCSVM bergantung pada beberapa faktor yang sangat mempengaruhi hasil klasifikasi seperti subset fitur yang digunakan untuk melatih model OCSVM, jenis *kernel*, dan *hyperparameternya*. Hasil eksperimen menunjukkan bahwa metode yang diusulkan mengungguli semua algoritma lainnya dalam hal *true positive rate* dan *false positive rate* untuk semua jenis perangkat IoT. Selain itu, ia mencapai waktu deteksi terendah, sekaligus mengurangi jumlah fitur yang dipilih secara signifikan.[22]

2.7 *Confusion Matrix*

Confusion Matrix merupakan salah satu *tools* dari *library sklearn* yang ampuh untuk melakukan penilaian kinerja dengan mengukur klasifikasi yang tumpang tindih. *Confusion matrix* bekerja dengan cara membuat tabel yang terdiri dari dua baris dan kolom seperti pada tabel 2.1 berikut.[23]

Tabel 2.1. Tabel Confusion Matrix

True Label			
	1	**TP**	**FN**
	0	**FP**	**TN**
		1	0
		Predicted Label	

Tabel 2.1 menunjukkan 2 label yaitu *True Label* sebagai nilai aktual dan *Predicted Label* sebagai hasil prediksi algoritma. Nilai 0 dan 1 merupakan isi dari label. Nilai 0 menunjukkan data yang normal, dan nilai 1 menunjukkan data anomali. Terdapat 4 klasifikasi label dari tabel 2.1 yaitu,

a. *True Positive (TP)*
   Nilai prediksi algoritma adalah 1 dan nilai aktual adalah 1. Ini menunjukkan jumlah prediksi algoritma terhadap nilai 1 secara tepat.

b. *False negative (FN)*

Nilai prediksi algoritma adalah 0 dan nilai aktual adalah 1. Ini menunjukkan jumlah prediksi algoritma terhadap nilai 1 yang tidak tepat.

c. *False Positive (FP)*
Nilai prediksi algoritma adalah 1 dan nilai aktual adalah 0. Ini menunjukkan jumlah prediksi algoritma terhadap nilai 0 yang tidak tepat.

d. *True Negative (TN)*
Nilai prediksi algoritma adalah 0 dan nilai aktual adalah 0. Ini menunjukkan jumlah prediksi algoritma terhadap nilai 0 secara tepat.

Klasifikasi tersebut berguna untuk mengukur tingkat kinerja algoritma yang telah dibangun dengan bantuan,

a. *Accuracy*
Menunjukkan proporsi dari prediksi yang benar terhadap total jumlah prediksi.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

b. Precision
Menunjukkan proporsi dari prediksi positif yang benar terhadap total jumlah prediksi positif, mengukur ketepatan atau akurasi dari prediksi positif.

$$Precision = \frac{TP}{TP + FP}$$

c. *Recall*
Menunjukkan proporsi dari prediksi positif yang benar terhadap total jumlah instance positif yang sebenarnya. Recall mengukur kemampuan model untuk menangkap semua instance positif. Recall = TP / (TP + FN)

$$Recall = \frac{TP}{TP + FN}$$

d. F-1 Score
Menggabungkan *precision* dan *recall* menjadi satu nilai tunggal dengan mengambil rata-rata harmonis dari keduanya demi memberikan keseimbangan antara *precision* dan *recall*.

$$F1\text{-}Score = \frac{2 * Recall * Precision}{Recall + Precision}$$