# BAB I PENDAHULUAN

Serangan siber adalah… Jenis-jenis serangan siber diantaranya adalah… (salah satunya adalah DDoS).

Serangan DDoS merupakan salah satu jenis ancaman yang paling merusak dan sulit diatasi dalam dunia Cyber Security. Serangan ini bertujuan untuk melumpuhkan layanan online dengan membanjiri target dengan traffic network yang sangat tinggi dari berbagai sumber yang tersebar di internet. Akibatnya, layanan menjadi tidak tersedia bagi pengguna yang sah, dan dapat menyebabkan kerugian finansial yang signifikan, kerusakan reputasi, dan gangguan operasional. Umumnya serangan DDoS dilakukan dengan mengirimkan permintaan dalam jumlah yang masif menggunakan spoofing IP atau IP palsu kepada situs target. Hal ini dilakukan agar situs yang diserang tidak dapat beroperasi secara optimal dikarenakan tidak dapat menangani berbagai permintaan hingga melampaui batas kemampuan dari sebuah situs [1].

Fenomena serangan DDoS terus berkembang baik dari segi frekuensi, intensitas, maupun kompleksitasnya. Hal ini menuntut adanya pemahaman yang mendalam dan pendekatan yang efektif untuk mendeteksi, mencegah, dan memitigasi dampak dari serangan tersebut. Dilansir dari halaman media digital IT “infosecurity” , jumlah kasus DDoS pada kuartal pertama tahun 2019 meningkat sebanyak 84% dari kuartal keempat tahun 2018 (Hill, 2019) [1].

Fakta lain 1

Fakta lain 2

Maka dari itu, diperlukan identifikasi serangan DDoS yang dilakukan untuk mencegah masalah–masalah diatas. Mengidentifikasi serangan DDoS menjadi tantangan yang lebih kompleks karena ada berbagai jenis strategi serangan DDoS. Beberapa jenis serangan DDoS yaitu ICMP flood, SYN flood, IP packet flood, dan lain-lain [2].

# BAB II PENELITIAN YANG TERKAIT

Klasifikasi dan prediksi serangan DDoS dapat dilakukan dengan pendekatan yang beragam. Seperti pada jurnal [3], para peneliti melakukan klasifikasi dan prediksi serangan DDoS menggunakan dataset UNSW-nb 15. Model yang digunakan dalam penelitian tersebut adalah Random Forest dan XGBoost Classifier. Hasilnya, masing-masing model mendapatkan skor akurasi rata-rata di angka 89% dan 90%.

Penelitian lain terkait deteksi serangan DDoS dilakukan oleh [2]. Para peneliti tersebut menggunakan dataset CICIDS2018 dan model naïve bayes dalam mengklasifikasi serangan DDoS. Eksperimen yang dilakukan menghasilkan performa akurasi 95%.

Peneliti pada [4] mengunakan dataset UNSW-nb15 dengan membagi data menjadi 70% sebagai train data dan 30% sebagai test data. Model yang digunakan yaitu naïve bayes, dengan menerapkan seleksi fitur information gain menggunakan 17 fitur berdasarkan nilai entrophy. Hasilnya, rata-rata akurasi yang didapat diangka 75.57%.

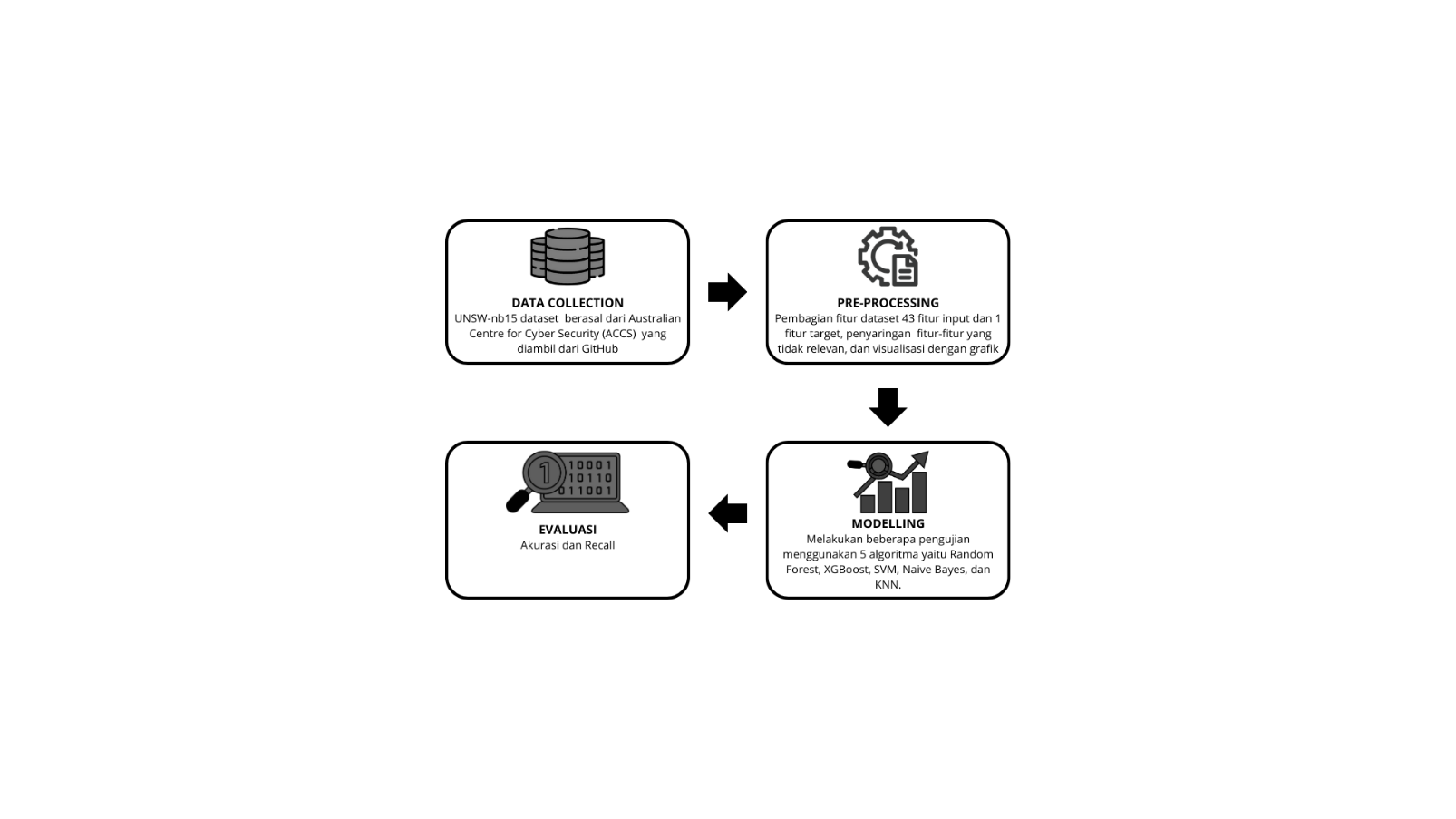
Peneliti pada [5] menggunakan dataset DDoS yang diperoleh dari situs Kaggle.com dan melakukan beberapa percobaan dengan beragam algoritma. Hasilnya, Logistic Regression, Random Forest, dan Naïve Bayes mendapatkan rata-rata akurasi sebesar 98,32%.

Penelitian pada [6] menggunakan dataset dari M. Teguh Kurniawan berupa data Disertasi serangan Ddos pada jaringan SDN. Dengan menggunakan Random Forest, skor akurasi yang diperoleh adalah sebesar 98%.

Berdasarkan beberapa penelitian tersebut, dibutuhkan penelitian lanjutan untuk meningkatkan performa dan mendapatkan hasil yang lebih baik dalam klasifikasi dan deteksi DDoS berbasis machine learning.

# BAB III METODE PENILITAN

Penelitian yang dilaksanakan ini adalah tentang deteksi serangan DDoS berbasis Machine Learning. Tahapan yang dilakukan pada penelitian ini dapat dilihat pada *Gambar 1*.



Gambar 1. Tahapan penelitian

## Hardware dan Software

Untuk mendukung kelancaran dan keberhasilan penelitian ini, dibutuhkan dukungan software dan hardware yang mumpuni. Hardware dengan spesifikasi yang tinggi, tidak akan bekerja optimal tanpa didukung software yang baik. Dan software yang berkualitas tidak akan bekerja secara baik jika tanpa dukungan hardware yang baik pula.

Software yang digunakan dalam penelitian ini, yaitu Orange Data Mining (<https://orangedatamining.com/>) atau yang biasa disebut dengan Orange. Software Orange memiliki widgets yang lengkap dan mudah digunakan dalam proyek data mining. Widget-widget tersebut berguna dalam kegiatan Preprocessing, Modeling, hingga Evaluation.

Sementara itu, Hardware yang digunakan dalam penelitian ini menggunakan sebuah *Personal Computer (PC)* dengan spesifikasi sebagai berikut:

* CPU : Intel Core i3 10100 @3.60 GHz
* RAM : 8 GB
* Storage : 256 GB (SSD)
* VGA : GeForce GTX 1650

## Data Collection

Dataset yang digunakan pada penelitian ini yaitu UNSW-NB15 dataset. Dataset ini merupakan dataset publik yang disediakan oleh Australian Centre for Cyber Security (ACCS) dan dapat diakses melalui GitHub repository. *Tabel 1* menunjukan informasi lengkap dari dataset.

**Tabel 1. Informasi Dataset UNSW-NB15**

|  |  |
| --- | --- |
| **Dataset** | UNSW\_NB15 |
| **Ukuran** | 175341 baris, 45 kolom |
| **Fitur** | 4 categorical, 39 numeric |
| **Target** | tipe categorical dengan 2 classes |
| **Metas** | 1 text |

## Pre-Processing

Pre-Processing merupakan kegiatan mengolah data mentah menjadi data yang siap diproses. Data mentah atau *raw data* seringkali memiliki masalah seperti ketidaklengkapan data, inkonsistensi data, kurang terstruktur, dan bahkan mengandung kesalahan nilai. Tujuannya dari tahap pre-processing ini adalah untuk mempersiapkan data dengan cermat agar lebih optimal digunakan dalam tahap berikutnya, yaitu modelling [5].

Dalam tahap pre-processing ini, 1 variabel dipilih sebagai variable target (*dependent*), yaitu ‘label’. Selanjutnya, terdapat 2 variabel yang dihapus, yaitu variable ‘attac\_cat’ dan ‘proto’. 42 variabel yang lain digunakan sebagai fitur atau variabel *independent*.

Langkah kedua dalam tahap pre-processing adalah menghapus fitur-fitur yang berjenis kategorikal. Pada dataset yang digunakan di penelitian ini, terdapat 2 fitur yang berjenis kategorikal dan akan dihapus, yaitu ‘state’ dan ‘service’. Sampai tahap ini, tersisa 40 fitur yang dapat diproses di tahap selanjutnya.

Langkah ketiga dalam penelitian ini adalah Normalisasi fitur, yaitu dengan mengubah rentang fitur numerik menjadi antara nilai 0 dan 1. Dengan demikian, maka seluruh fitur pada dataset yang digunakan pada penelitian ini akan memiliki rentang yang sama.

Langkah ke empat, yaitu seleksi fitur. Pada penelitian ini, terdapat 2 metode seleksi fitur yang digunakan, yaitu Information Gain dan FCBC. 2 metode seleksi fitur berbarsis perankingan tersebut, berguna untuk mengurangi jumlah fitur sehingga performa algoritma machine learning dalam mendeteksi serangan DDoS dapat dioptimalkan.

## Modelling

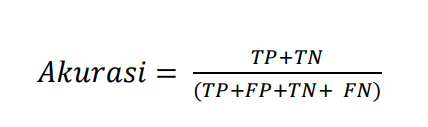
Pada tahap ini, beberapa algoritma machine learning diaplikasikan dan dianalisis untuk melihat performa masing-masing algoritma dalam mendeteksi serangan DDoS. Lima algoritma yang digunakan adalah: Random Forest, XGBoost, Support Vector Machine (SVM), Naive Bayes, dan K-Nearest Neighbors (KNN). Model-model ini dipilih karena keunggulan mereka dalam berbagai tugas klasifikasi dan popularitas mereka dalam penelitian machine learning.

## E. Evaluasi

Tahap evaluasi dalam penelitian ini melibatkan penghitungan akurasi dan recall untuk mengevaluasi kinerja model. Kedua metrik ini dipilih karena memberikan gambaran yang jelas tentang performa model dalam mengklasifikasikan data dengan benar.

1. Akurasi

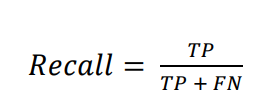
Akurasi berfungsi untuk mengukur seberapa sering model memprediksi dengan tepat. Nilai akurasi dihitung dengan membagi jumlah prediksi benar (*True Positive* dan *True Negative*) dengan total prediksi (*True Positive*, *False Positive, True Negative,* dan *False Negative*):



Akurasi digunakan untuk mengukur seberapa tepat model dalam pengklasifikasian sample secara keseluruhan.

1. F1-Score
2. Presisi
3. Recall

Recall mengukur seberapa baik model mengidentifikasi kategori penting. Ini dihitung dengan membagi jumlah True Positive dengan jumlah True Positive dan False Negative:



Dalam penilitian ini, recall digunakan untuk memastikan model tidak melewatkan banyak sampel yang termasuk dalam kategori penting [7].

# BAB IV HASIL DAN PEMBAHASAN

# REFERENSI

[1] M. K. Harto and A. Basuki, “Deteksi Serangan DDoS pada Jaringan Berbasis SDN dengan Klasifikasi Random Forest,” *J. Pengemb. Teknol. Inf. Dan Ilmu Komput.*, vol. 5, no. 4, Art. no. 4, Apr. 2021.

[2] M. Zidane, “Klasifikasi Serangan Distributed Denial-of-Service (DDoS) menggunakan Metode Data Mining Naive Bayes,” *J. Pengemb. Teknol. Inf. Dan Ilmu Komput.*, vol. 6, no. 1, Art. no. 1, 2022.

[3] Ismail *et al.*, “A Machine Learning-Based Classification and Prediction Technique for DDoS Attacks,” *IEEE Access*, vol. 10, pp. 21443–21454, 2022, doi: 10.1109/ACCESS.2022.3152577.

[4] I. B. G. Amartya, “Implementasi Algoritma Naive Bayes Classifier (NBC) Dan Information Gain Untuk Mendeteksi DDoS,” *J. Elektron. Ilmu Komput. Udayana*, vol. 11, no. 2, pp. 273–282, Nov. 2022.

[5] S. Joses, S. Quinevera, R. Mardianto, D. Yulvida, and A. M. Shiddiqi, “Pendekatan Metode Ensemble Learning untuk Deteksi Serangan DDoS menggunakan Soft Voting Classifier,” *JEPIN J. Edukasi Dan Penelit. Inform.*, vol. 10, no. 1, pp. 79–87, Apr. 2024, doi: 10.26418/jp.v10i1.73241.

[6] A. Ekawijana, A. Bakhrun, and M. T. Kurniawan, “Deteksi Serangan DDOS Pada Jaringan SDN dengan Metode Random Forest,” *J. MEDIA Inform. BUDIDARMA*, vol. 8, no. 1, Art. no. 1, Jan. 2024, doi: 10.30865/mib.v8i1.6928.

[7] “Deteksi Malware menggunakan Metode Stacking berbasis Ensemble | Rafrastara | Jurnal Informatika: Jurnal Pengembangan IT.” Accessed: Jun. 24, 2024. [Online]. Available: http://ejournal.poltekharber.ac.id/index.php/informatika/article/view/4606