BAB 1
PENDAHULUAN

1.1. Latar Belakang

Malware adalah perangkat lunak yang diciptakan dengan tujuan tertentu, dimana penyerang mencari celah keamanan dalam sistem. Dampak buruk dari *Malware* dapat dirasakan pada komputer atau pengguna, karena penyerang dapat mencuri informasi atau data pribadi seseorang. (Adenansi & Novarina, 2017) Tujuan pembuatan *Malware* oleh penyerang dapat bervariasi, termasuk merusak atau membobol suatu sistem operasi melalui skrip rahasia, dan dapat diinisiasi secara tersembunyi. *Malware* memiliki kemampuan untuk menginfiltrasi sistem operasi, mengakibatkan sistem komputer menggunakan sumber daya tanpa pengetahuan pemilik perangkat, bahkan mengeksploitasi untuk mengumpulkan informasi pribadi yang nantinya dibagikan kepada pihak ketiga tanpa persetujuan pengguna. Beberapa varian klasik *Malware*, seperti *adware, spyware, ransomware*, berbagai jenis virus (*overwriting virus, prepending virus, appending virus, file infector virus, boot sector virus, multipartie virus, dan macro virus), Worms*, dan *Trojan Horse* (*remote access trojan, password sending trojan, keylogger, destructive trojan, FTP trojan, software detection killer, procy trojan*), dapat mengancam keamanan dan menurunkan performa jaringan.(Damanik dkk., 2023)

Implementasi sistem deteksi berbasis web memungkinkan pemantauan aktif terhadap aktivitas dan aliran data. Apabila sistem mendeteksi pola atau perilaku yang mencurigakan, seperti perubahan tiba-tiba dalam pola komunikasi atau upaya akses yang tidak sah, sistem dapat memberikan peringatan kepada pengguna atau bahkan mengambil tindakan otomatis, seperti mengisolasi perangkat yang terinfeksi. Pembuatan sistem deteksi ini melibatkan penerapan *deep learning*, yang dapat dijelaskan sebagai model jaringan yang terdiri dari sejumlah lapisan. Setiap lapisan tersebut terdiri dari simpul-simpul yang melakukan perhitungan (Rizki dkk., 2020). Pada simpul *input*, data dikalikan dengan bobot khusus, lalu hasilnya melewati fungsi aktivasi untuk menentukan seberapa jauh sinyal akan bergerak melalui jaringan, yang nantinya memengaruhi hasil akhir.

Menurut penelitian yang dilakukan oleh (Kusumaningrum, 2018), *Convolution Neural Network (CNN)* merupakan salah satu algoritma yang berasal dari *Deep learning*, yang merupakan pengembangan dari *Multi Layer Perceptron* (MLP). *Convolutional Neural Network* (*CNN*) adalah jenis arsitektur jaringan saraf buatan yang secara khusus dirancang untuk menangani pengolahan data *grid*, seperti gambar atau data visual. *CNN* memiliki kemampuan untuk efektif mengekstrak fitur-fitur penting dari data visual, menjadikannya sangat sesuai untuk tugas-tugas seperti klasifikasi gambar, deteksi objek, dan segmentasi gambar. Arsitektur *CNN* juga dapat diterapkan dengan baik dalam konteks klasifikasi dataset *Malware*. Meskipun awalnya dirancang untuk tugas pengolahan gambar, *CNN* telah berhasil digunakan dalam berbagai tugas klasifikasi, termasuk klasifikasi dataset *Malware*, seperti yang dibuktikan dalam penelitian (Sahu dkk., 2021) yang menggabungkan *CNN* dengan *LSTM* dan mencapai tingkat akurasi sebesar 96%.

Penelitian ini memiliki tujuan untuk mengembangkan sebuah sistem deteksi *Malware* berbasis web dengan menggunakan dataset IoT23, yang telah dikembangkan oleh (Garcia dkk., 2020). Dataset ini terdiri dari 23 rekaman atau skenario yang mencakup berbagai lalu lintas jaringan *IoT*. Skenario-skenario tersebut dibagi menjadi dua puluh tangkapan jaringan (*file pcap*) yang berasal dari perangkat IoT yang terinfeksi, yang masing-masing memiliki sampel *Malware* dieksekusi, dan tiga tangkapan jaringan dari lalu lintas jaringan perangkat IoT yang tidak terinfeksi.

Sebuah penelitian sebelumnya yang dilakukan oleh (Liang & Vankayalapati, 2021) menggunakan dataset yang sama, menerapkan metode *Convolutional Neural Network (CNN*) yang menghasilkan akurasi sebesar 69%, *Decision Tree* sebesar 73%, *SVM* 69%, dan *Naïve Bayes* sebesar 30%. Dalam konteks penelitian ini, sebuah upaya akan dilakukan untuk membuat sebuah aplikasi berbasis web yang mendeteksi serangan-serangan *Malware*.

Berdasarkan latar belakang yang sudah diuraikan, dapat disimpulkan bahwa judul penelitian ini adalah "KLASIFIKASI *MALWARE* MENGGUNAKAN METODE *CONVOLUTIONAL NEURAL NETWORK* (*CNN*) BERBASIS WEBSITE"

1.2. Rumusan Masalah

Berdasarkan latar belakang yang sudah dijelaskan, rumusan masalah pada penelitian ini adalah:

1. Bagaimana cara membuat aplikasi berbasis web yang mendeteksi serangan *Malware* menggunakan metode *CNN*?
2. Bagaimana akurasi, presisi, *recall* dan *f1-score* yang dihasilkan oleh metode *CNN?*

1.3. Tujuan Penelitian

Tujuan dilakukannya penelitian ini adalah sebagai berikut:

1. Membuat sebuah aplikasi berbasis web untuk mendeteksi *Malware* menggunakan metode CNN
2. Mengetahui akurasi, presisi, *recall*, dan *f1-score* yang dihasilkan oleh metode *CNN* dalam aplikasi tersebut.

1.4. Manfaat Penelitian

Adapun manfaat yang bisa didapat dari penelitian ini yaitu:

1. Memberikan kontribusi dalam pengembangan teknologi keamanan *cyber* dengan menyediakan aplikasi berbasis web yang mampu mendeteksi serangan *malware* menggunakan metode CNN.
2. Memberikan wawasan mendalam mengenai penerapan metode *deep learning* khususnya CNN dalam mendeteksi *malware.*
3. Menyediakan data empiris yang penting untuk mengevaluasi kinerja metode CNN dalam konteksi deteksi *malware.*

1.5. Batasan Penelitian

Adapun batasan-batasan penelitian yang ditetapkan adalah sebagai berikut:

1. Penelitian membuat sebuah aplikasi berbasis web untuk memprediksi serangan *Malware* berdasarkan data latih yang diberikan kepada model.
2. Algoritma yang digunakan adalah *Convolutional Neural Network (CNN).*
3. Dataset acuan pada penelitian ini adalah dataset IoT23 oleh (Garcia dkk., 2020).

4. *Environment* yang digunakan adalah *Jupyter Notebook* untuk *preprocessing* data dan bahasa pemrograman *Python* sebagai Bahasa pemrograman untuk membangun model klasifikasi.
5. *Framework* yang digunakan untuk membangun aplikasi berbasis web pada penelitian ini adalah *Flask.*