BAB I
PENDAHULUAN

Pada bab ini akan dijelaskan tentang latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian dan ruang lingkup dari penelitian yang dilakukan.

1.1 Latar Belakang

Saat ini masyarakat Indonesia sangat memperdebatkan masalah keamanan siber, Salah satu contohnya adalah serangan siber yang terjadi pada Pusat Data Nasional (PDN) Kementerian Komunikasi dan Informatika (Kemenkominfo) pada Kamis, 20 Juni 2024. Beberapa layanan publik, termasuk layanan imigrasi telah terganggu karena serangan terhadap fasilitas pengelolaan data oleh beberapa lembaga pemerintah [1]. Maraknya kasus peretasan yang dialami oleh beberapa server Pemerintah merupakan suatu fenomena yang tidak bisa terhindarkan, karena perkembangan ilmu pengetahuan dan teknologi di kawasan global yang sangat pesat. Institusi Pendidikan seperti pada Universitas XYZ menjadi salah satu institusi yang harus diselamatkan dari maraknya kasus peretasan karena mereka menyimpan dan mengelola berbagai jenis data sensitif, seperti data pribadi mahasiswa, staf, dan informasi akademik yang mana data tersebut menjadi data yang harus dilindungi dari serangan siber.

Serangan siber yang umum terjadi meliputi *phishing*, di mana penyerang menipu korban untuk memberikan informasi sensitif melalui email atau situs web palsu. *Malware* seperti virus, *ransomware*, dan *trojan* yang merusak sistem atau mencuri data. *Distributed Denial of Service (DDoS)* yang membanjiri *server* dengan lalu lintas berlebihan untuk menghentikan akses layanan. *SQL injection* yang mengeksploitasi kelemahan aplikasi web untuk mengambil atau mengubah data di dalam *database. Brute force attack* yang mencoba menebak kata sandi melalui percobaan berbagai kombinasi. *Man in the middle (MitM)* yang menyusup ke dalam komunikasi antar pengguna untuk mencuri atau memanipulasi data. Serta *social engineering*, di mana penyerang memanipulasi individu untuk memberikan informasi sensitif [2].

Permasalahan serangan siber diperparah dengan minimnya tenaga ahli di bidang keamanan siber yang membuat deteksi dan respons terhadap serangan menjadi tidak optimal dan mengakibatkan kerugian besar baik dari sisi data maupun finansial. Keterbatasan sistem deteksi yang ada tidak mampu untuk mengidentifikasi ancaman yang belum pernah terlihat, sulitnya membedakan antara lalu lintas jaringan yang sah dengan aktivitas mencurigakan, serta kurangnya integrasi dan pembaruan secara *realtime* pada sistem deteksi yang digunakan. Selain itu, serangan yang menggunakan teknik *social engineering* sering kali tidak dapat dideteksi oleh perangkat lunak, karena serangan tersebut menargetkan kelemahan manusia dari pada sistem.

Penelitian terdahulu telah mengkaji berbagai metode dalam mendeteksi serangan siber, di antaranya penelitian oleh Radford et al. (2019) menunjukkan bahwa arsitektur *Transformer* seperti BERT dan RoBERTa memiliki kemampuan yang lebih baik dalam memahami konteks dari data berbasis teks, yang dapat diterapkan dalam klasifikasi serangan siber berbasis log sistem [28]. Selain itu, penelitian oleh Vaswani et al. (2017) yang memperkenalkan *Transformer* sebagai model pemrosesan sekuensial berbasis *self-attention* telah menjadi dasar bagi berbagai model modern dalam analisis data berbasis teks [3]. Di sisi lain, Howard & Ruder (2018) memperkenalkan Universal *Language* Model Fine-Tuning (ULMFiT), yang menunjukkan efektivitas *transfer learning* dalam meningkatkan performa model *deep learning* untuk berbagai tugas klasifikasi teks [4].

Dengan mempertimbangkan penelitian terdahulu, pendekatan menggunakan metode *Transformer* dan optimasi melalui *transfer learning* terbukti menjadi solusi yang lebih efektif dibandingkan metode tradisional seperti *Decision Tree, Support Vector Machine* (SVM), dan *Naive Bayes* dalam klasifikasi serangan siber. Model *Transformer*, seperti BERT dan RoBERTa, memiliki kemampuan memahami konteks dan pola dalam data teks kompleks, serta lebih unggul dalam menangani data sekuensial melalui mekanisme self-attention [11]. Hal ini memungkinkan model mendeteksi berbagai jenis serangan secara lebih akurat dan cepat dibandingkan metode konvensional yang cenderung gagal mengenali pola tidak terstruktur pada data *log* jaringan.

Pada penelitian Alshamrani et al. (2020) menunjukkan bahwa *Transformer* yang di-*fine-tune* untuk klasifikasi serangan siber mampu mengidentifikasi anomali dalam data *log* dengan akurasi lebih dari 90%, bahkan pada jenis serangan baru (*zero-day attack*) [12]. Studi kasus serupa diterapkan dalam lingkungan pendidikan, di mana sistem informasi akademik rentan terhadap serangan seperti SQL *Injection* dan DDoS. *Log server* mencatat pola anomali berupa permintaan POST dengan *query string* mencurigakan dan lonjakan trafik dalam hitungan detik. Penerapan model BERT yang dioptimasi dengan *transfer learning* terhadap *log* tersebut terbukti meningkatkan kemampuan klasifikasi serangan hingga 92% akurasi dibanding *baseline model tradisional* [16].

Transfer learning sendiri memungkinkan pemanfaatan model pra-latih dari korpus besar seperti CIC-IDS2017 atau NSL-KDD, yang kemudian diadaptasi pada domain server lokal. Teknik ini tidak hanya mengurangi waktu pelatihan, tetapi juga meningkatkan generalisasi terhadap data baru, termasuk pola serangan yang belum pernah dilabeli secara manual [18]. Dengan demikian, penerapan metode *Transformer* dan *transfer learning* berkontribusi signifikan dalam memperkuat sistem pertahanan siber, terutama dalam konteks institusi pendidikan yang menyimpan data sensitif mahasiswa dan staf.

1.2 Rumusan Masalah

Dalam penelitian ini diuraikan beberapa rumusan masalah yang akan menjadi fokus penelitiannya adalah sebagai berikut:

1. Bagaimana mengklasifikasikan jenis serangan siber yang terjadi pada *server* dengan menggunakan metode *Transformer* BERT dan RoBERTa serta *transfer learning* ULMFit?
2. Bagaimana cara menyiapkan data serangan siber (data *preprocessing*) agar dapat digunakan untuk melatih model *Transformer* BERT dan RoBERTa serta *transfer learning* ULMFit?
3. Apakah metode *transfer learning* ULMFit dapat mengoptimasi metode *Transformer* BERT dan RoBERTa dalam klasifikasi serangan siber?

4. Bagaimana cara mengevaluasi performa model *Transformer* BERT dan RoBERTa serta *transfer learning* ULMFit dalam mengklasifikasikan serangan siber?

1.3 Tujuan Penelitian

Berdasarkan dari beberapa rumusan masalah yang diuraikan tersebut, maka dapat kami tentukan tujuan dari penelitian ini adalah sebagai berikut:

1. Mengembangkan sistem otomatis untuk mengklasifikasikan jenis serangan siber yang terjadi pada *server*, sehingga dapat meningkatkan efisiensi dan akurasi dalam mendeteksi ancaman siber.
2. Mempersiapkan data serangan siber (data *preprocessing*) dengan metode yang tepat agar dapat digunakan sebagai *dataset* dalam melatih model *Transformer* BERT dan RoBERTa serta *transfer learning* ULMFit, sehingga model dapat mengenali pola serangan dengan lebih baik.
3. Menganalisis efektivitas metode *transfer learning* ULMFit dalam meningkatkan kinerja model *Transformer* BERT dan RoBERTa dalam klasifikasi serangan siber, dengan membandingkan hasilnya terhadap model *Transformer* yang dilatih dari awal.
4. Mengevaluasi performa model *Transformer* BERT dan RoBERTa serta *transfer learning* ULMFit dalam mengklasifikasikan serangan siber berdasarkan metrik evaluasi seperti akurasi, presisi, *recall*, dan F1-*score*, guna memastikan kehandalan model dalam mendeteksi ancaman di lingkungan *server* Universitas XYZ.

1.4 Manfaat Penelitian

Penelitian "Optimasi Klasifikasi Serangan Siber Pada *Server* Menggunakan Metode *Transformer* Dan *Transfer Learning*" memiliki berbagai manfaat, baik secara akademik maupun praktis. Berikut adalah uraian manfaat dari penelitian ini:

1. Penerapan metode *Transformer* BERT dan RoBERTa serta *transfer learning* ULMFit untuk klasifikasi serangan siber yang dapat mendeteksi serangan dengan lebih akurat dan cepat.
2. Optimasi dalam proses klasifikasi serangan siber dengan metode *transfer learning* ULMFit ini mempercepat proses klasifikasi sehingga meningkatkan efisiensi waktu pemprosesan.
3. Memberikan informasi perbandingan kinerja sistem deteksi serangan siber antara metode *Transformer* BERT dan RoBERTa serta *transfer learning* ULMFit.
4. Memberikan manfaat bagi peningkatan keamanan siber dengan mengimplementasikan model deteksi serangan yang telah dikembangkan, pihak universitas dapat mendeteksi pola serangan dari *log* HTTP, mengklasifikasikan jenis ancaman secara otomatis, dan meresponsnya secara cepat sebelum berdampak luas.

1.5 Batasan Masalah

Dalam upaya memperjelas ruang lingkup penelitian sehingga lebih fokus dan terarah maka kami uraikan batasan masalah pada penelitian ini adalah sebagai berikut:

1. Penelitian ini akan mengklasifikasikan jenis serangan siber DDoS dan SQL *Injection*.
2. *Dataset* yang digunakan dalam penelitian ini berasal dari *dataset* yang dikumpulkan dari serangan jaringan pada *server* Universitas XYZ.
3. Penelitian ini menggunakan model *Transformer* BERT dan RoBERTa serta *transfer learning* ULMFit.
4. Evaluasi performa model klasifikasi dilakukan dengan menggunakan metrik *accuracy*, *precision*, *recall*, dan F1-*score.*
5. Pengimplementasian model *Transformer* BERT dan RoBERTa akan dioptimasi dengan model *transfer learning* ULMFit.
6. Eksperimen pada penelitian yang mencakup *traning* dan pengujian dilakukan dalam lingkungan simulasi.

1.6 Ruang Lingkup

Penelitian ini dilakukan dengan mengumpulkan serangan jaringan pada *server* Universitas XYZ dengan parameter *Wireshark, CPU, RAM, Disk*. Fokus penelitian ini pada upaya mengklasifikasikan serangan siber pada jaringan *DDoS* dan serangan *SQL injection*, karena keduanya secara statistik merupakan ancaman yang paling umum dan kritis terhadap infrastruktur layanan web. Menurut laporan dari OWASP (*Open Web Application Security Project*), SQL *Injection* secara konsisten masuk dalam daftar *Top* 10 kerentanan keamanan aplikasi web, karena sifatnya yang dapat mengekspos data sensitif dan memanipulasi basis data melalui input yang tidak tervalidasi [32]. Serangan ini sangat berbahaya karena dapat dilakukan dengan mudah melalui antarmuka pengguna tanpa membutuhkan akses internal. Di sisi lain, DDoS adalah jenis serangan volumetrik yang sangat merusak ketersediaan layanan, dengan melumpuhkan server atau jaringan melalui lonjakan trafik yang berlebihan. Menurut laporan dari Cisco dan Cloudflare, DDoS menjadi salah satu metode serangan yang paling banyak digunakan untuk melumpuhkan layanan perusahaan dan pemerintah, dan insidennya meningkat seiring dengan adopsi layanan daring [33]. Dengan demikian penelitian ini akan memberikan solusi tepat guna bagi Universitas XYZ dalam mencegah kerusakan data atau gangguan operasional yang diakibatkan oleh serangan siber. Metode *transformer* yang digunakan untuk mengklasifikasikan serangan siber dengan algoritma *BERT* dan *RoBERTa* serta akan di optimalkan dengan model *transfer learning* sehingga dapat meningkatkan performanya. Penelitian ini hanya mencakup pengklasifikasian serangan siber pada tingkat pola lalu lintas jaringan (*network traffic*). Penelitian tidak akan mencakup deteksi serangan pada level aplikasi spesifik atau teknik mitigasi serangan.