BAB I
PENDAHULUAN

Bab pendahuluan ini membahas mengenai latar belakang masalah, rumusan masalah, tujuan dan manfaat, ruang lingkup, serta sistematika penulisan yang akan digunakan dalam dokumen skripsi ini yang berjudul Optimasi *Anomaly Based Intrusion Detection System* Menggunakan Arsitektur *Multilayer Perceptron* Dengan Metode Seleksi Fitur *Analysis Of Variance*.

1.1 Latar Belakang

Perkembangan teknologi internet yang pesat dalam beberapa dekade terakhir telah membawa perubahan mendasar dalam cara manusia berinteraksi dengan dunia di sekitarnya (Lone dkk., 2023). Salah satu manifestasi paling signifikan dari transformasi ini adalah kemunculan *Internet of Things* (IoT), sebuah paradigma teknologi yang memungkinkan miliaran perangkat fisik untuk terhubung, berkomunikasi, dan bertukar data secara *real-time* melalui jaringan digital tanpa memerlukan intervensi manusia (Alam & Ansari, 2022). Kemampuan ini membuka peluang luar biasa di berbagai sektor strategis kehidupan modern, mulai dari layanan kesehatan cerdas (*smart healthcare*) yang memungkinkan pemantauan kondisi pasien secara jarak jauh, rumah cerdas (*smart homes*) yang mengotomatisasi pengelolaan lingkungan tempat tinggal, transportasi cerdas (*smart transportation*) yang mengoptimalkan mobilitas dan keselamatan berkendara, hingga pendidikan cerdas (*smart education*) yang memperluas akses dan personalisasi pembelajaran (Al-Garadi dkk., 2020). Dengan proyeksi jumlah perangkat IoT yang terus tumbuh secara eksponensial, ekosistem ini menjanjikan tingkat efisiensi dan otomatisasi yang sebelumnya tidak pernah terbayangkan (Hassija dkk., 2019).

Namun, di balik segala manfaat yang ditawarkan, perkembangan ekosistem IoT yang masif justru membuka celah keamanan yang serius dan mengkhawatirkan. Mayoritas perangkat IoT dirancang dengan mengutamakan fungsionalitas dan efisiensi biaya, sehingga kerap beroperasi dengan sumber daya komputasi yang sangat terbatas serta menggunakan protokol keamanan yang minimal (Alam & Ansari, 2022). Kondisi ini menjadikan perangkat-perangkat tersebut sebagai target yang sangat rentan bagi berbagai bentuk serangan siber. Salah satu ancaman yang paling merusak adalah serangan *Distributed Denial*

of Service (DDoS), di mana ribuan hingga jutaan perangkat IoT yang telah berhasil dikompromikan dikerahkan secara bersamaan untuk membanjiri suatu target dengan permintaan dalam jumlah masif dan waktu singkat (Zikria dkk., 2020). Akibatnya, server target mengalami kelebihan kapasitas dan tidak mampu melayani lalu lintas yang sah dari pengguna yang sebenarnya (Rohit dkk., 2019). Dampak dari serangan semacam ini sangat nyata dan berbahaya, terutama ketika menyasar layanan-layanan kritis seperti sistem pemantauan kesehatan pasien, infrastruktur rumah cerdas, hingga sistem kendali transportasi, di mana gangguan layanan dapat berujung pada kerugian besar bahkan mengancam keselamatan jiwa (Mohy-eddine dkk., 2024).

Sebagai respons awal terhadap ancaman tersebut, komunitas keamanan siber mengembangkan dan mengadopsi *Signature-based Intrusion Detection System* (SIDS) sebagai mekanisme pertahanan konvensional yang telah lama diandalkan. SIDS bekerja dengan prinsip pencocokan pola (*pattern matching*), yaitu membandingkan setiap lalu lintas jaringan yang masuk dengan basis data yang berisi *signature* atau tanda tangan dari serangan-serangan yang telah diketahui sebelumnya (Ravindran dkk., 2025). Pendekatan ini terbukti sangat efektif dan mampu menghasilkan akurasi yang tinggi selama serangan yang dihadapi termasuk dalam jenis yang telah terdokumentasikan dalam basis data tersebut (Khraisat dkk, 2019). Sayangnya, efektivitas SIDS memiliki batas yang sangat jelas. Ketergantungannya pada pembaruan basis data secara manual menjadi kelemahan fundamental yang tidak dapat diatasi dalam ekosistem IoT yang dinamis dan terus berkembang. SIDS sama sekali tidak mampu mendeteksi serangan varian baru, serangan polimorfik, maupun serangan *zero-day*, yaitu serangan yang memanfaatkan celah keamanan yang belum pernah diketahui atau terdokumentasikan sebelumnya, karena pola serangan tersebut memang belum tersedia dalam basis datanya (Buczak & Guven, 2016). Dengan lanskap ancaman siber yang terus berevolusi dengan sangat cepat, keterbatasan mendasar ini membuat SIDS tidak lagi memadai sebagai satu-satunya lini pertahanan.

Menyadari keterbatasan kritis SIDS, perhatian penelitian dan industri keamanan siber beralih kepada *Anomaly-based Intrusion Detection System* (AIDS) atau sistem deteksi intrusi berbasis anomali sebagai solusi yang lebih adaptif dan relevan (Jyothsna & Prasad, 2020). Di mana anomali disini mengacu pada AIDS yang bekerja dengan membangun profil perilaku normal dari lalu lintas jaringan, kemudian mengidentifikasi setiap aktivitas yang menyimpang atau anomali secara signifikan dari profil tersebut sebagai potensi serangan

(Hassan & Daneshwar, 2023). Dengan cara kerja ini, AIDS secara inheren mampu mendeteksi serangan *zero-day* karena tidak bergantung pada keberadaan signature yang telah diketahui sebelumnya. AIDS akan memicu peringatan ketika suatu perilaku dinilai berbeda dari pola normalnya, tanpa perlu mengenali jenis serangan tersebut secara spesifik. (Khraisat dkk., 2019) Karakteristik inilah yang menjadikan AIDS sebagai pendekatan yang jauh lebih relevan dan tangguh untuk menghadapi ancaman siber modern di lingkungan IoT yang terus berubah.

Untuk memaksimalkan kemampuan AIDS, penelitian lanjutan secara intensif mengintegrasikan teknik *Machine Learning* (ML) dan *Deep Learning* (DL) ke dalam sistem deteksi anomali. Pada fase awal, algoritma ML klasik seperti *K-Nearest Neighbors* (KNN), *Support Vector Machine* (SVM), dan *Random Forest* (RF) banyak digunakan karena keunggulannya dalam hal interpretabilitas, efisiensi komputasi, dan kemudahan implementasi. Azimjonov & Kim (2024) menunjukkan bahwa *Linear* SVM pada dataset intrusi seperti NSL-KDD mampu mencapai akurasi hingga 99,78%, sementara Luqman dkk. (2025) membuktikan konsistensi metode *ensemble* RF dan SVM pada dataset intrusi serupa yaitu UNSW-NB15. Meskipun demikian, performa algoritma ML klasik cenderung menurun ketika dihadapkan pada volume data yang masif, berdimensi tinggi, dan mengandung pola yang sangat kompleks seperti pada dataset Bot-IoT. Keterbatasan ini mendorong adopsi teknik yang menawarkan kemampuan ekstraksi fitur secara otomatis dan pemodelan hubungan non-linear yang jauh lebih unggul (García-Teodoro dkk., 2009).

Berbagai arsitektur DL telah dieksplorasi dan terbukti memberikan performa superior dalam konteks deteksi intrusi pada lingkungan IoT yang kompleks dan dinamis. Koroniotis dkk. (2019) membuktikan bahwa metode berbasis sekuensial seperti *Long Short-Term Memory* (LSTM) mampu mencapai akurasi sebesar 99,74%, sementara *Recurrent Neural Network* (RNN) mencapai 99,55% pada dataset Bot-IoT, yang menunjukkan kemampuannya dalam menangkap pola temporal dari lalu lintas jaringan. Di sisi lain, Roopak dkk. (2019) mengusulkan arsitektur hybrid CNN+LSTM yang mengombinasikan kemampuan ekstraksi fitur spasial dari *Convolutional Neural Network* (CNN) dengan pemodelan dependensi waktu dari LSTM, dan berhasil mencapai akurasi sebesar 97,16% khususnya dalam mendeteksi serangan DDoS. Hasil-hasil tersebut menunjukkan bahwa pemilihan arsitektur yang tepat, baik berbasis sekuensial maupun *hybrid*, sangat berpengaruh

terhadap performa sistem deteksi intrusi, terutama dalam menghadapi karakteristik data IoT yang bersifat kompleks, besar, dan beragam.

Meskipun demikian, dataset intrusi seperti NSL-KDD, UNSW-NB15, dan Bot-IoT pada dasarnya merupakan data tabular terstruktur berupa fitur-fitur statistik hasil ekstraksi lalu lintas jaringan (Chimphlee & Chimphlee, 2023). Karakteristik ini menjadikan asumsi dependensi spasial yang melandasi CNN maupun dependensi temporal yang dieksploitasi LSTM sebagian besar tidak relevan dan bahkan berpotensi merugikan karena memperkenalkan kompleksitas yang tidak dibutuhkan di mana pada lingkungan seperti IoT, kompleksitas model merupakan hal yang krusial (Airlangga, 2024). Dalam konteks inilah *Multilayer Perceptron* (MLP) menunjukkan keunggulannya sebagai arsitektur yang memetakan hubungan antar fitur secara langsung tanpa asumsi struktural yang tidak perlu. Hal ini didukung oleh beberapa temuan bahwa MLP menghasilkan performa yang setara dengan CNN dan LSTM dengan nilai *f1-score* yang hampir identik (Ali dkk., 2025), juga unggul dibandingkan pendekatan ML klasik seperti RF, SVM, dan AdaBoost (Kanimozhi & Jacob, 2019a), dan mencapai akurasi 99,97% pada dataset CSE-CIC-IDS2018 (Kanimozhi & Jacob, 2019b).

Meski menawarkan akurasi yang tinggi, penerapan model *deep learning* pada dataset IoT berskala besar juga menimbulkan tantangan utama yang tidak dapat diabaikan. Permasalahan seperti beban komputasi yang sangat tinggi menjadi isu krusial, di mana model-model kompleks membutuhkan waktu pelatihan yang panjang, konsumsi memori yang besar, serta dukungan infrastruktur GPU yang mahal (Rafique dkk., 2024). Kondisi ini membuat implementasi model menjadi kurang realistis, terutama pada perangkat IoT yang umumnya memiliki keterbatasan sumber daya komputasi dan energi. Selain itu, kompleksitas model juga dapat menyebabkan peningkatan latensi saat proses inferensi, yang berpotensi menghambat deteksi serangan secara real-time. Oleh karena itu, diperlukan pendekatan yang lebih efisien, seperti optimasi arsitektur atau reduksi dimensi fitur, agar performa tinggi tetap dapat dicapai tanpa mengorbankan efisiensi sistem (Logeswari dkk., 2025).

Untuk menjawab tantangan tersebut, penerapan teknik seleksi fitur (*feature selection*) sebelum proses pelatihan model menjadi langkah tepat. Seleksi fitur bertujuan untuk mengidentifikasi dan mempertahankan hanya fitur-fitur yang paling relevan dan

diskriminatif, sekaligus membuang fitur-fitur yang redundan atau tidak informatif, Sehingga dapat mengurangi kompleksitas model yang berpengaruh kepada beban komputasi (J. Li dkk., 2018). Mohy-eddine dkk. (2023) membuktikan bahwa reduksi fitur yang efektif mampu menurunkan waktu pelatihan hingga 41,2%, bahkan meningkatkan *f1-score* dari 97,27% menjadi 99,59% menggunakan *Whale Optimization Algorithm* (WOA). Penelitian milik Musthafa dkk. (2024) juga membktikan keunggulan dari seleksi fitur yang dapat mengurangi ukuran model dan *loading time* menggunakan metode seleksi fitur *Spearmann Correlation* dan *Analysis of Variance* (ANOVA), di mana ANOVA menunjukkan performa yang lebih superior dibanding *Spearmann*. ANOVA berhasil mengurangi ukuran model menjadi hanya 1278 Kb dari 1314 Kb dengan peningkatan akurasi menjadi 99,73%. X. Liu dkk. (2021) juga membuktikan penggunaan ANOVA dapat mengurangi kompleksitas dimensi, dan bahkan meningkatkan performa model. Dari hal tersebut tersebut, ANOVA dinilai paling efektif dan sesuai untuk karakteristik dataset Bot-IoT yang bersifat numerik, berskala besar, dan berfokus pada klasifikasi multikelas, karena kemampuannya dalam mengukur signifikansi statistik perbedaan antar kelas serangan secara langsung.

Berdasarkan uraian di atas, penelitian ini mengusulkan pendekatan yang mengintegrasikan metode seleksi fitur berbasis ANOVA dengan arsitektur MLP untuk optimisasi sistem deteksi intrusi berbasis anomali pada jaringan IoT menggunakan dataset Bot-IoT dalam skenario klasifikasi multikelas jenis serangan. Pendekatan ini diharapkan menghasilkan model yang tidak hanya unggul secara akurasi, tetapi juga efisien secara komputasi sehingga dapat memberikan kontribusi nyata bagi pengembangan sistem keamanan IoT yang lebih *robust*, transparan, dan dapat diandalkan.

1.2 Rumusan Masalah

Rumusan masalah penelitian ini adalah bagaimana efektivitas dan performa model *deep learning* berbasis arsitektur MLP dalam mendeteksi intrusi multikelas pada dataset Bot-IoT dengan integrasi seleksi fitur menggunakan metode ANOVA untuk mereduksi dimensi data, serta apakah pendekatan tersebut lebih efektif dibandingkan dengan metode seleksi fitur lainnya.

1.3 Tujuan dan Manfaat Penelitian

Tujuan penelitian ini adalah mengembangkan dan mengevaluasi performa model deteksi intrusi berbasis arsitektur MLP pada dataset Bot-IoT, dengan mengintegrasikan

seleksi fitur metode ANOVA untuk mengidentifikasi fitur yang paling diskriminatif dan mereduksi dimensi data. Selain itu, penelitian ini juga bertujuan membandingkan performa model dari sisi akurasi dan efisiensi komputasi dengan penelitian lain.

Manfaat yang diharapkan adalah tersedianya model klasifikasi multikelas yang andal dalam mengidentifikasi jenis serangan dengan akurasi tinggi, serta diperolehnya subset fitur signifikan yang dapat digunakan untuk menyederhanakan sistem deteksi intrusi berbasis signature (SIDS). Hasil ini juga diharapkan mendukung pengembangan sistem keamanan IoT yang lebih efisien, transparan, dan adaptif, serta menjadi referensi untuk penelitian lanjutan.

1.4 Ruang Lingkup Penelitian

Ruang lingkup penelitian ini meliputi pengembangan dan evaluasi model deteksi anomali jaringan dengan fokus pada dataset BoT-IoT dari UNSW-NB15. Proses penelitian dibatasi pada tahap prapemrosesan data, seleksi fitur berbasis korelasi, pelatihan model *deep learning*, dan evaluasi performa menggunakan metrik seperti akurasi, *presicison*, *recall*, dan *F1-score* dan perbandingan dengan model dasar. Penelitian ini tidak mencakup implementasi langsung pada sistem IDS *real-time* maupun optimasi arsitektur jaringan komputer secara fisik.

1.5 Sistematika Penulisan

Struktur penulisan dalam laporan ini dibagi menjadi lima bab utama yang membahas penelitian berjudul "Optimasi *Anomaly Based Intrusion Detection System* Menggunakan Arsitektur *Multilayer Perceptron* Dengan Metode Seleksi Fitur *Analysis Of Variance*". Pembagian ini disusun dengan tujuan agar penulisan menjadi lebih sistematis dan mudah dipahami. Berikut ini adalah ringkasan singkat dari setiap bab yang dibahas:

BAB I PENDAHULUAN

Bab ini berisi latar belakang masalah terkait deteksi intrusi pada IoT, rumusan masalah, tujuan dan manfaat penelitian, ruang lingkup, serta sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Bab ini menguraikan teori pendukung seperti IoT, *intrusion detection system* (IDS), *deep learning* (MLP), serta metode seleksi fitur ANOVA, dan penelitian terdahulu yang relevan.

BAB III METODOLOGI

Bab ini menjelaskan tahapan penelitian, meliputi *preprocessing* data, seleksi fitur dengan ANOVA, perancangan model MLP, serta proses pelatihan dan evaluasi.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil seleksi fitur dan performa model, serta analisis efektivitas metode yang digunakan.

BAB V KESIMPULAN

Bab ini berisi kesimpulan penelitian dan saran untuk pengembangan selanjutnya.