BAB I
PENDAHULUAN

1.1 Latar Belakang

Di era digital modern, ancaman keamanan jaringan semakin kompleks seiring meningkatnya volume dan keragaman trafik data. Sistem deteksi intrusi tradisional berbasis *signature* sering kali gagal mengenali serangan *zero-day* atau pola serangan baru yang belum terdaftar dalam basis data. Sebagai solusi, pendekatan *unsupervised anomaly* detection seperti *Isolation Forest* menjadi relevan karena kemampuannya mendeteksi anomali tanpa memerlukan label data. Nugroho et al., (2025) mengimplementasikan *Isolation Forest* pada dataset CIC-IDS2017 dan menunjukkan performa deteksi anomali yang baik dengan *false positive* rate rendah. Keunggulan algoritma ini dalam memproses data log juga dikonfirmasi oleh Santoso & Wahyuni, (2024), yang berhasil mendeteksi aktivitas anomali pada log web server secara efektif.

Namun, penggunaan *Isolation Forest* secara tunggal memiliki keterbatasan, terutama terkait bias pada pembagian ruang fitur yang hanya mengikuti sumbu koordinat. Varian seperti *Extended Isolation Forest (EIF)* dikembangkan untuk mengatasi hal ini dengan menggunakan hiperplane miring, di mana Sunarto et al., (2021) mendemonstrasikan bahwa *EIF* lebih stabil dan akurat dalam mendeteksi *outlier*. Selain masalah bias, tantangan utama dalam trafik jaringan adalah memahami struktur data yang sangat heterogen. Di sinilah peran teknik klasterisasi menjadi krusial. Feng et al., (2022) menunjukkan bahwa penggunaan X-means (turunan K-Means) sangat relevan untuk mengklasifikasikan trafik normal dan anomali berdasarkan kemiripan pola karakteristiknya.

Alasan utama penggabungan *Isolation Forest* dan *K-Means* dalam penelitian ini didasarkan pada kebutuhan akan sistem yang tidak hanya mampu mengisolasi titik data asing, tetapi juga memahami struktur kelompok trafik yang kompleks. Laskar et al., (2021) mengusulkan model hibrida yang menggabungkan *K-Means* dan *Isolation Forest* untuk deteksi anomali skala besar *(big data),* yang terbukti sukses melakukan deteksi *real-time* pada trafik industri. Kombinasi ini bekerja secara

sinergis: *K-Means* berfungsi untuk memetakan kepadatan dan profil perilaku trafik yang serupa, sementara *Isolation Forest* secara efisien memisahkan anomali yang muncul di luar profil kelompok tersebut.

Implementasi *Isolation Forest* juga telah teruji dalam mendeteksi ancaman internal yang sulit dikenali oleh sistem biasa, sebagaimana ditunjukkan oleh Palaniappan et al., (2025) dalam analisis *insider threat*. Dengan adanya riset terbaru seperti OptIForest yang mengoptimasi struktur pohon (Xiang et al., 2023), terlihat bahwa pengembangan algoritma berbasis isolasi masih menjadi fokus utama dalam keamanan siber. Berdasarkan tinjauan literatur tersebut, penerapan kombinasi *Isolation Forest* dan *K-Means* menjadi sangat tepat untuk penelitian ini. Integrasi tersebut menggabungkan kekuatan isolasi (deteksi *outlier*) dengan struktur cluster (pemahaman pola trafik), sehingga memungkinkan terciptanya sistem deteksi anomali yang adaptif dan akurat terhadap dinamika trafik jaringan yang kompleks.

1.2 Rumusan Masalah

Berdasarkan uraian pada bagian latar belakang, terlihat bahwa sistem keamanan jaringan konvensional sering kali tidak mampu beradaptasi dengan dinamika trafik yang masif dan gagal mendeteksi ancaman yang belum terdaftar dalam basis data *signature*. Penggunaan metode tunggal sering kali belum optimal, *Isolation Forest* secara mandiri masih memiliki kerentanan terhadap bias ruang fitur, sementara klasterisasi murni seperti *K-Means* tidak dirancang secara khusus untuk memisahkan *outlier* secara efisien. Ketidakmampuan untuk menggabungkan pemahaman struktur trafik dengan kecepatan isolasi anomali menjadi hambatan dalam menciptakan sistem pertahanan yang tangguh. Permasalahan utama dalam penelitian ini dirumuskan sebagai berikut:

1. Sistem deteksi intrusi tradisional belum mampu mengatasi serangan *zero-day* karena ketergantungan pada basis data *signature* yang statis, sehingga pola serangan baru sering kali tidak terdeteksi.
2. Penggunaan algoritma *Isolation Forest* secara tunggal masih tidak terlepas dari keterbatasan bias pada pembagian ruang fitur koordinat, yang menyebabkan

proses isolasi titik anomali pada trafik yang sangat heterogen menjadi kurang akurat.

3. Algoritma klasterisasi seperti *K-Means* secara mandiri tidak dirancang untuk mendeteksi *outlier*, sehingga gagal memisahkan anomali secara efisien jika tidak dikombinasikan dengan metode isolasi yang tepat.

1.3 Tujuan Penelitian

Tujuan penelitian ini disusun sebagai upaya untuk mengatasi berbagai kendala dalam deteksi keamanan jaringan yang telah diidentifikasi pada rumusan masalah. Secara umum, penelitian ini bertujuan untuk membangun model deteksi anomali yang lebih adaptif dan akurat dengan memanfaatkan sinergi antara teknik isolasi dan klasterisasi. Secara khusus, tujuan penelitian ini adalah sebagai berikut:

1. Mengatasi ketidakmampuan sistem konvensional dalam mengenali ancaman baru dengan mengimplementasikan algoritma *Isolation Forest* yang mampu mendeteksi anomali tanpa ketergantungan pada basis data *signature*.
2. Meminimalkan bias pemisahan data pada fitur yang kompleks dengan memanfaatkan algoritma K-*Means* untuk memetakan struktur kelompok trafik, sehingga identifikasi anomali menjadi lebih presisi.
3. Menghilangkan celah kelemahan pada penggunaan algoritma tunggal melalui penggabungan *Isolation Forest* dan *K-Means* guna meningkatkan performa deteksi pada trafik jaringan yang heterogeny.

Dengan tercapainya tujuan-tujuan tersebut, penelitian ini diharapkan dapat menghasilkan sebuah model deteksi anomali yang tidak hanya akurat, tetapi juga dapat diimplementasikan sebagai solusi pendukung dalam meningkatkan keamanan jaringan. Selain itu, hasil penelitian ini juga diharapkan dapat menjadi referensi dalam pengembangan metode deteksi ancaman berbasis *machine learning* pada penelitian selanjutnya.

1.4 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan kontribusi baik secara teoritis maupun praktis dalam bidang keamanan jaringan dan penerapan metode deteksi anomali berbasis *machine learning.* Manfaat penelitian ini meliputi:

1. Manfaat Teoritis

   Penelitian ini dapat menambah wawasan dan pengetahuan mengenai penerapan algoritma *unsupervised learning*, khususnya Isolation Forest dan K-Means, dalam mendeteksi anomali pada trafik jaringan. Selain itu, penelitian ini dapat menjadi referensi bagi pengembangan kajian akademik yang berfokus pada teknik deteksi intrusi berbasis pembelajaran mesin.

2. Manfaat Praktis

   Hasil penelitian ini dapat dimanfaatkan oleh perusahaan atau organisasi untuk meningkatkan keamanan jaringan dengan memanfaatkan model deteksi anomali yang adaptif dan tidak bergantung pada *signature*. Model yang dihasilkan juga dapat dijadikan dasar untuk mengembangkan aplikasi sederhana sebagai alat pendukung pemantauan trafik jaringan secara otomatis.

3. Manfaat bagi Penelitian Selanjutnya

   Penelitian ini dapat menjadi acuan bagi peneliti berikutnya yang ingin mengembangkan metode *hybrid* dalam deteksi anomali, baik dengan menambahkan algoritma lain, memperluas jenis data, atau menerapkannya pada lingkungan jaringan yang berbeda. Dengan demikian, penelitian ini berpotensi menjadi pijakan awal untuk pengembangan solusi keamanan jaringan yang lebih komprehensif.

1.5 Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen untuk menganalisis dan menguji efektivitas integrasi algoritma *machine learning* dalam mendeteksi anomali pada trafik jaringan. Metode ini dipilih untuk memberikan bukti empiris mengenai sejauh mana penggabungan dua algoritma mampu menutupi kelemahan sistem deteksi tradisional. Tahapan penelitian disusun secara sistematis sebagai berikut:

1. Pengumpulan dan Pra-Pemrosesan Data

   Tahap awal dimulai dengan memperoleh dataset trafik jaringan yang mencakup berbagai profil aktivitas, baik normal maupun anomali. Dataset tersebut melalui proses *data cleaning* dan normalisasi fitur untuk memastikan data siap diolah, mengingat trafik jaringan memiliki karakteristik yang heterogen dan dimensi yang tinggi.

2. Implementasi Model Integrasi

   Pada tahap inti, dilakukan penerapan algoritma *K-Means* dan *Isolation Forest* secara sinergis. *K-Means* digunakan terlebih dahulu untuk memetakan struktur kepadatan trafik ke dalam klaster-klaster tertentu guna memahami profil perilaku normal. Selanjutnya, *Isolation Forest* diterapkan untuk mengisolasi titik data yang muncul di luar profil kelompok tersebut, guna mengatasi keterbatasan bias pada pembagian fitur yang sering ditemukan pada penggunaan algoritma secara tunggal.

3. Pengujian dan Evaluasi Kinerja

   Tahap ini melibatkan pengujian model untuk menilai kemampuan sistem dalam mendeteksi anomali pada berbagai skenario trafik. Evaluasi dilakukan menggunakan metrik performa seperti *Accuracy*, *Precision*, *Recall*, dan *False Positive Rate (FPR)*. Hal ini bertujuan untuk memastikan bahwa model yang dikembangkan benar-benar mampu meminimalkan kegagalan identifikasi anomali dan bekerja secara akurat terhadap dinamika trafik jaringan yang kompleks.

4. Analisis dan Penarikan Kesimpulan

   Analisis terhadap hasil evaluasi menunjukkan efektivitas kolaborasi kedua metode ini. Poin-poin inilah yang nantinya menjadi landasan untuk menciptakan pertahanan jaringan yang lebih kuat guna menangkal serangan siber yang terus bermunculan.