BAB III
METODE PENELITIAN

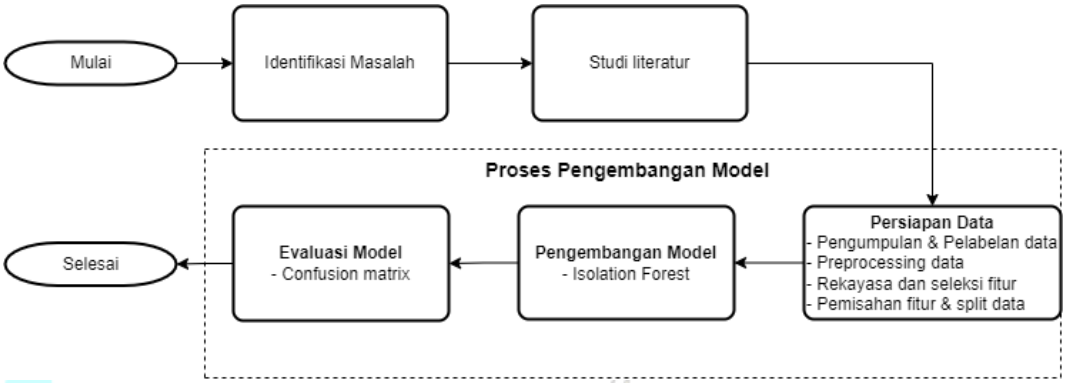
3.1. Objek Penelitian

Penelitian ini berfokus pada pemodelan deteksi anomali pada jaringan *IP Camera* menggunakan algoritma *Isolation Forest* serta pengujian kinerja model yang dihasilkan. Model tersebut digunakan untuk mendeteksi secara dini adanya gangguan pada jaringan *IP Camera* tertentu, seperti kamera *offline*, latensi tinggi, atau kehilangan paket yang signifikan. Dataset yang digunakan dalam penelitian ini terdiri dari data log ping jaringan *IP Camera*. Data ini dikumpulkan melalui proses ping secara berkala untuk memantau kondisi jaringan setiap kamera. Hasil yang diharapakan dari penelitian adalah model deteksi anomali yang mampu mengidentifikasi apakah data tersebut termasuk kategori normal atau anomali berdasarkan pola-pola data, seperti *Round Trip Time*, *Response Rate*, tingkat *Packet loss* dan lain-lain. Model ini dirancang untuk diterapkan pada sistem pemantauan jaringan yang dapat memberikan peringatan dini melalui notifikasi *real-time*. Metode ini diharapakan dapat mengatasi keterbatasan dalam mendeteksi gangguan yang selama ini masih bergantung pada laporan dari operator, sehingga dapat mendukung upaya pemeliharaan yang lebih responsif dan efektif.

Penelitian ini dilaksanakan di Perum Peruri, tepatnya pada Departemen Pengamanan, Seksi Pengamanan Elektronik. Seksi Pengamanan Elektronik memiliki tanggung jawab utama dalam memastikan keamanan melalui teknologi elektronik, termasuk dengan CCTV atau *IP Camera*. Sistem keamanan elektronik ini berperan penting dalam mendukung operasional perusahaan, terutama dalam memantau aktivitas di area strategis dan melindungi aset penting perusahaan. Adapun waktu pelaksanaan penelitian ini berlangsung selama 4 bulan dari November 2024 hingga Februari 2025.

3.2. Prosedur Penelitian

Tahapan penelitian tersebut dapat dilihat pada Gambar 3.1, yang menjelaskan proses penelitian mulai dari identifikasi masalah, studi literatur, persiapan data, hingga pengembangan dan evaluasi model *Isolation Forest.*



Gambar 3.1 Alur Penelitian

3.2.1 Identifikasi Masalah

Tahap pertama dimulai dengan mengidentifikasi permasalahan terkait deteksi anomali pada jaringan *IP Camera*. Permasalahan ini mencakup keterbatasan metode pemantauan tradisional yang masih bergantung pada laporan manual dari operator kepada tim pemeliharaan, terutama ketika terdapat kamera yang *offline* atau mengalami gangguan pada tampilan visual. Metode ini bersifat reaktif dan memakan waktu, terutama ketika jumlah perangkat yang di pantau sangat banyak dan jumlah personel terbatas. Akibatnya, kamera yang mengalami gangguan sering kali terlambat terdeteksi atau bahkan terlewatkan. Permasalahan ini kemudian menjadi dasar untuk mengembangkan solusi berbasis *Machine Learning* menggunakan algoritma *Isolation Forest*, yang di rancang untuk mendeteksi secara dini adanya gangguan pada jaringan *IP Camera*.

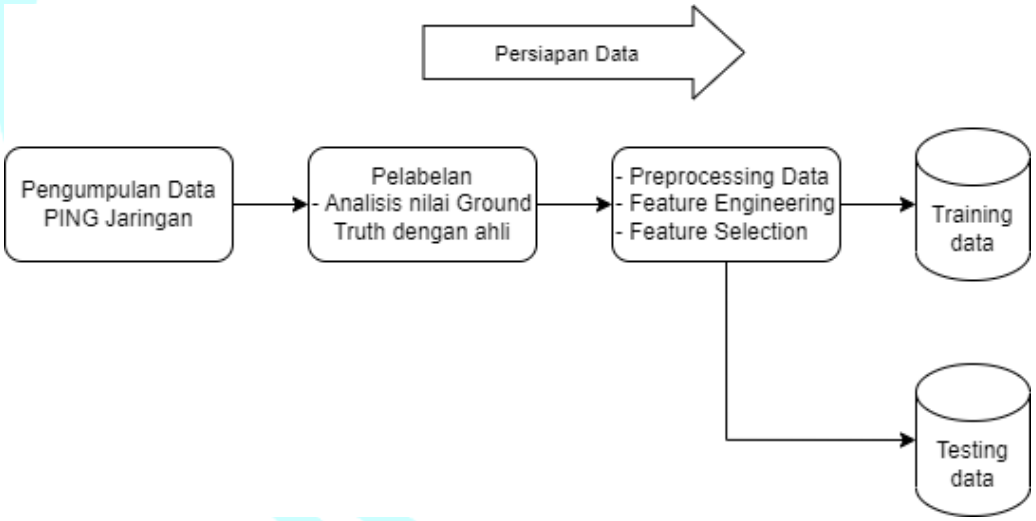
3.2.2 Studi Literatur

Pada tahap ini, dilakukan pencarian bahan pustaka yang relevan untuk mendukung landasan teori dan konteks penelitian. Proses ini mencakup pengumpulan artikel, jurnal, buku, dan sumber informasi lainnya yang berkaitan dengan topik penelitian. Sumber-sumber tersebut bertujuan untuk memberikan wawasan mendalam mengenai teori, metode, dan hasil penelitian sebelumnya,

sehingga membantu peneliti memahami perkembangan isu yang diteliti, mengidentifikasi kesenjangan penelitian, serta merumuskan kerangka konsep yang tepat.

Hasil penelusuran literatur dalam penelitian ini mencakup 24 artikel ilmiah yang berkaitan dengan penggunaan model *Isolation Forest* dalam deteksi anomali. Selain itu, informasi tambahan juga diperoleh melalui diskusi bersama tim ahli yang memahami kondisi anomali pada jaringan *IP Camera* berdasarkan log hasil PING, sehingga memperkuat landasan dalam perumusan masalah dan pendekatan penelitian.

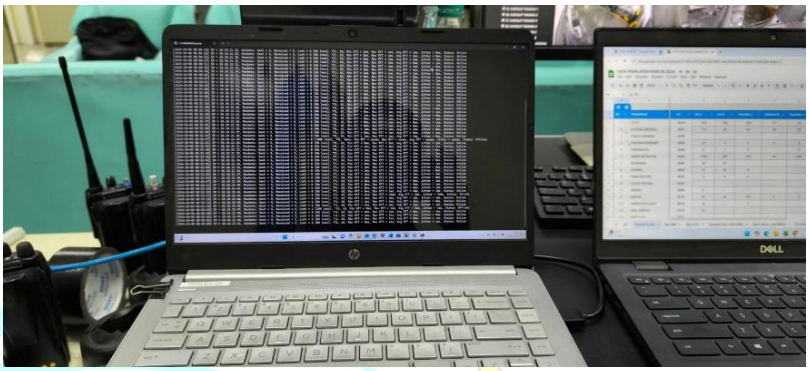
3.2.3 Persiapan Data



Gambar 3.2 Tahap Persiapan Data

Tahap persiapan data (*Data Preparation*) mencakup proses pengumpulan data, pelabelan data, *preprocessing*, *feature engineering*, *feature selection* dan data *splitting*. Data yang digunakan pada penelitian ini merupakan data log PING jaringan *IP Camera*. Pengumpulan data dilakukan dengan cara melakukan PING secara berkala terhadap 100 IP kamera yang terhubung ke dalam jaringan. Data yang dikumpulkan mencakup *timestamp*, *IP Address*, *Packet Sent*, *Packet Received*, *Time to Live (TTL)*, *Round Trip Time (RTT)*, *Jitter*, dan *Execution Time*. Data tersebut merupakan data primer, karena diperoleh langsung oleh peneliti melalui proses PING terhadap *IP Camera* dalam jaringan. Proses pengumpulan dilakukan

secara otomatis dengan interval waktu tertentu untuk memastikan data yang diperoleh mencerminkan berbagai kondisi jaringan.



Gambar 3.3 Dokumentasi Proses Pengumpulan Data.

Proses PING dilakukan menggunakan bahasa pemrograman *Python*, dan hasilnya disimpan dalam format CSV untuk mempermudah pengolahan data lebih lanjut. Data yang terkumpul kemudian akan melalui tahap pelabelan atau analisis nilai *Ground Truth* yaitu proses menentukan label atau nilai referensi yang benar untuk suatu data, yang nantinya digunakan sebagai dasar evaluasi dalam menilai perfoma model algoritma *Isolation Forest*. Proses analisis dan pelabelan data ini dilakukan dengan melibatkan ahli jaringan, tim IT, atau senior pemeliharaan sistem yang berpengalaman dalam pemantauan jaringan dan memahami karakteristik data yang tergolong anomali. Contoh Sampel data hasil PING pada jaringan *IP Camera* dapat dilihat pada Tabel 3.1.

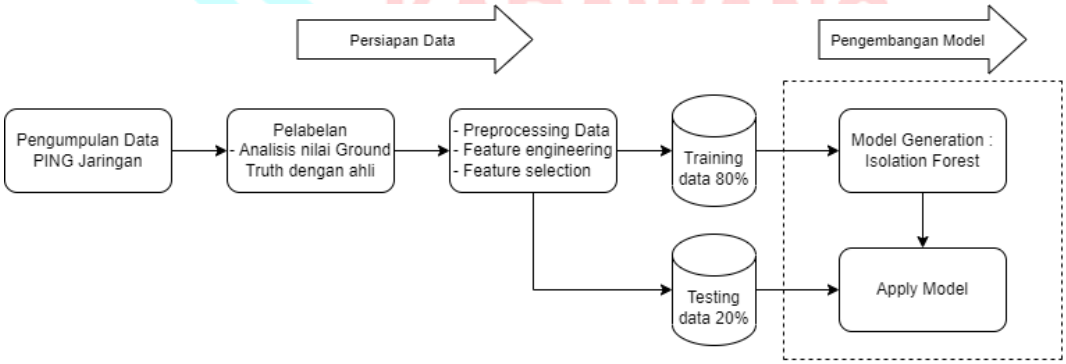
Tabel 3. 1 Sampel Dataset

Timestamp	IP Address	Packets Sent	Packets Received	Packets Lost
2024-11-20 13:48:26	10.10.3.100	5	5	0
2024-11-20 13:48:29	10.10.3.2	5	2	3
2024-11-20 13:48:33	10.10.1.22	5	5	0
2024-11-20 13:48:36	10.10.1.4	5	5	0
2024-11-20 13:48:39	10.10.1.12	5	5	0

TTL	Min RTT (ms)	Max RTT (ms)	Avg RTT (ms)	Jitter (ms)	Time (s)
64	6	17	10	4	4,096
63	2	72	37	70	15,801
64	15	29	18	4,25	4,113
64	6	14	8	3,25	4,098
64	14	36	20	6,50	4,122

Setelah proses pelabelan data, dataset melalui tahap *Preprocessing* untuk meningkatkan kualitas data, dan mempersiapkan data untuk proses analisis atau pelatihan pada model deteksi anomali *Isolation Forest* (Risky Pratiwi et al., 2024). Tahap *Preprocessing* data mencakup *Data Cleaning*, dan *Data Transformation*, setelah itu dilanjutkan tahap *Feature Engineering*, *Feature Selection*, Pemisahan Fitur, dan Split data. Proses pembersihan data (*Data Cleaning*) dimulai dengan identifikasi dan penanganan nilai yang hilang (*Missing Value*) serta penghapusan data duplikat guna meningkatkan kebersihan dan akurasi data. Selanjutnya, dilakukan *Data Transformation*, yang mencakup *label encoding* untuk mengonversi data kategorikal menjadi format numerik serta penyesuaian tipe data. Pada tahap *feature engineering*, dilakukan rekayasa fitur untuk menghasilkan fitur baru yang lebih informatif dari data yang tersedia, seperti RTT *Range*, dan Log RTT. Setelah itu dilakukan seleksi fitur untuk memilih subset fitur yang paling relevan dan berkontribusi untuk deteksi anomali. Setelah data siap digunakan, dilakukan proses pemisahan fitur menjadi dua bagian yaitu Fitur X dan Label Y, serta data akan di bagi menjadi 80% sebagai data pelatihan (*training data*) dan 20% sebagai data pengujian (*testing data*) pada tahap split data.

3.2.4 Pengembangan Model



Gambar 3.4 Tahap Pengembangan Model

Pada tahap pengembangan model, algoritma *Isolation Forest* dilatih dengan membagi dataset menjadi 80% data latih (*training data*) dan 20% data uji (*testing data*). Meskipun *Isolation Forest* merupakan Algoritma *Unsupervised Learning* yang secara prinsip tidak memerlukan pembagian data, dalam penelitian ini, pembagian data dilakukan dengan tujuan untuk mengevaluasi dan memvalidasi

kinerja model dalam mendeteksi anomali pada data yang belum dilihat sebelumnya, meskipun model tersebut tidak menggunakan label saat pelatihan. Pendekatan ini memungkinkan pengukuran perfoma model dengan membandingkan hasil prediksi terhadap G*round Truth* pada data uji, menggunakan metrik evaluasi seperti Akurasi, *Precision, Recall,* dan *F1-Score*.

Model dibangun menggunakan data latih untuk menemukan pola yang dapat membedakan anomali dari data normal. Proses ini melibatkan pembentukan pohon-pohon isolasi, dimana setiap pohon bekerja dengan membagi data secara berulang hingga terisolasi sepenuhnya. Hasil pelatihan ini, model akan menghasilkan nilai skor anomali (*anomaly score*) untuk setiap data. Nilai ini digunakan untuk menentukan apakah data termasuk anomali atau tidak. Berikut adalah tahapan-tahapan dari algoritma *Isolation Forest*, dapat dilihat pada *Algorithm* 2 :

Algorithm 2 iForest

Input:

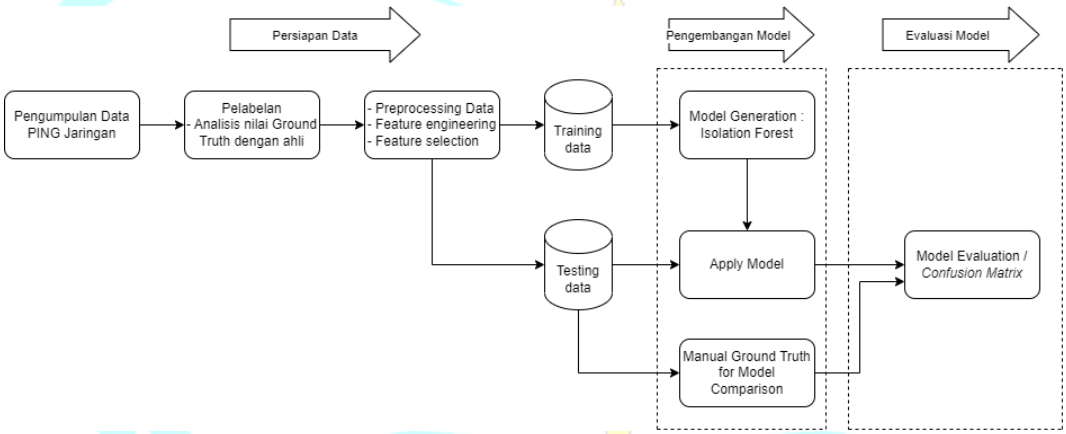
- X : Data Latih, Data Uji
- M : Fitur
- T : Jumlah iTree
- Q : Threshold

Output: R_X : Hasil klasifikasi (Anomali / Normal)

1: Bangun Isolation Forest dengan membuat T Isolation Trees :
2: Inialisasi $X' \leftarrow X$
3: **While** $X' \neq c$ **do :**
4: Pilih fitur secara acak $m \leftarrow randomFeature(M)$
5: Pilih titik split secara acak $q \leftarrow randomSplit(m)$
6: Bagi subset X' berdasarkan q dan m
7: **End**
8: Hitung panjang jalur rata-rata $\bar{h}(x)$ untuk setiap instance x dengan persamaan 2.1
9: Hitung skor anomali $s(x, n)$ untuk setiap instance x berdasarkan panjang jalur rata-rata dengan persamaan 2.2
10: Tentukan apakah instance x merupakan anomali atau normal:
11: **if** $s(x, n) \geq$ Q **maka :**
12: $R_X =$ Anomali
13: **else:**
14: $R_X =$ Normal
15: **end if**

Setelah model selesai tahap pelatihan, selanjutnya model diterapkan pada data uji untuk mengukur kemampuannya dalam mendeteksi anomali pada data yang belum pernah digunakan sebelumnya. Model akan memberikan skor anomali untuk setiap data dalam set pengujian dan mengklasifikasikanya sebagai normal atau anomali berdasarkan *threshold* tertentu. Hasil pengujian ini kemudian digunakan untuk menilai sejauh mana model dapat mengidentifikasi anomali dengan baik sebelum dilakukan analisis lebih lanjut pada tahap evaluasi model.

3.2.5 Evaluasi Model



Gambar 3.5 Tahap Evaluasi Model

Setelah dilakukan pengujian, model di evaluasi untuk mengukur perfomanya dalam mendeteksi anomali. Evaluasi ini dilakukan menggunakan beberapa metrik, yaitu precision, recall, F1-score, PR AUC (Precision-Recall – Area Under Curve), dan ROC AUC (Receiver Operating Characteristic – Area Under Curve). Precision mengukur sejauh mana model dapat menghindari kesalahan dalam mendeteksi anomali, yaitu rasio antara jumlah prediksi anomali yang benar terhadap seluruh prediksi anomali. Recall menunjukan kemampuan model dalam menemukan seluruh anomali yang sebenarnya ada dalam dataset. F1-score merupakan kombinasi precision dan recall dalam satu metrik yang memberikan keseimbangan antara keduanya. Untuk mendukung evaluasi ini, digunakan confusion matrix, yaitu tabel yang menampilkan jumlah prediksi benar dan salah dari model sebagai dasar perhitungan metrik evaluasi seperti precision, recall, dan akurasi (Hikmayanti et al., 2023). Berikut adalah tabel confusion matrix di tunjukan pada Tabel 3.2. Tabel 3.2 *Confusion Matrix*

		Class Prediction	
		1	**0**
Actual class	**1**	*True positive* (TP)	*False positive* (FP)
	0	*False negative* (FN)	*True negative* (TN)

$$Akurasi = \frac{TP + TN}{TP + FP + FN + TN} \quad (3.1)$$

$$Precision = \frac{TP}{TP + FP} \quad (3.2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3.3)$$

$$F1 - score = \frac{Precision\ .Recall}{Precision + Recall} \quad (3.4)$$

Selain itu, ROC AUC dan PR AUC menunjukan kemampuan model dalam membedakan antara kelas normal dan anomali secara keseluruhan. Evaluasi pada set pengujian ini didasarkan pada label yang ada dalam dataset uji. Dari hasil evaluasi ini, dapat diperoleh gambaran mengenai efektivitas model dalam mendeteksi anomali serta membantu dalam menentukan apakah model yang digunakan sudah optimal atau masih memerlukan perbaikan lebih lanjut.