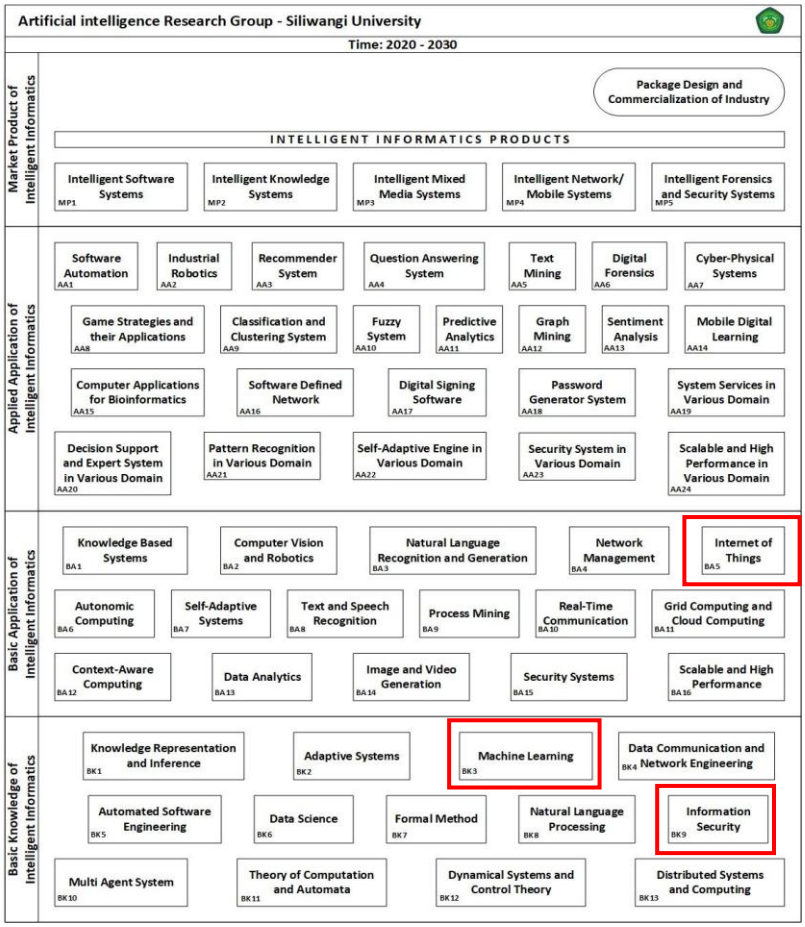
BAB III

METODOLOGI

3.1 Peta Jalan (*Road Map*) Penelitian

Penelitian yang diusulkan dalam laporan ini merupakan bagian dari *roadmap* penelitian KK ISI (Informatika dan Sistem Intelijen) di Universitas Siliwangi. Gambar 3.1 menggambarkan kajian penelitian yang akan dilakukan, dengan fokus pada area yang ditandai dalam garis persegi panjang berwarna merah, yaitu terkait *Internet of Things* (BA5), *Machine Learning* (BK3), dan *Information Security* (BK9).

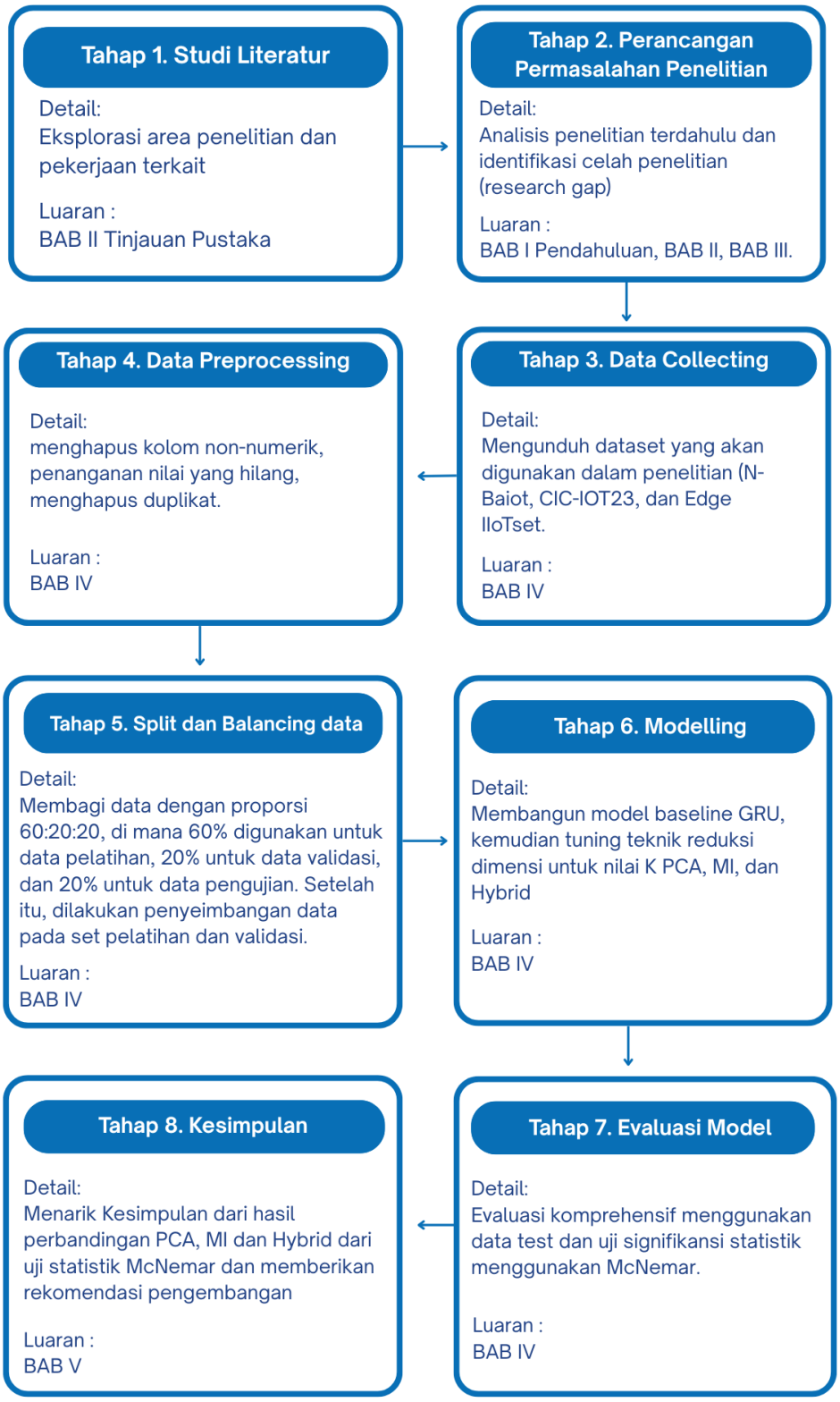


Gambar 3. 1 Roadmap Penelitian(AIS Universitas Siliwangi, 2020)

3.2 Tahapan Penelitian

Penelitian ini bertujuan untuk mengevaluasi efektivitas teknik reduksi dimensi dalam meningkatkan kinerja deteksi malware pada lingkungan IoT yang memiliki karakteristik data berdimensi tinggi (*high-dimensionality*). Proses diawali dengan penerapan tiga strategi reduksi dimensi, yaitu *Principal Component Analysis* (PCA), *Mutual Information* (MI), dan pendekatan *Hybrid* (MI+PCA), yang bertujuan untuk menyederhanakan kompleksitas data trafik jaringan tanpa menghilangkan informasi krusial.

Data yang telah diproses kemudian dilatih menggunakan arsitektur *Gated Recurrent Unit* (GRU). Kinerja model dengan reduksi dimensi ini selanjutnya akan dianalisis komparatif terhadap model *baseline* (tanpa reduksi dimensi) untuk mengukur peningkatan performa yang dihasilkan. Untuk menjamin keabsahan hasil, signifikansi perbedaan performa antar model divalidasi menggunakan uji statistik McNemar. Seluruh eksperimen dilakukan pada tiga dataset dengan karakteristik beragam, yaitu N-BaIoT, CIC-IoT2023, dan Edge-IIoTset, guna memastikan generalisasi metode. Adapun detail tahapan penelitian ini diilustrasikan pada Gambar 3.2.



Gambar 3. 2 Tahapan Penelitian

3.2.1 Studi Literatur

Tahap ini bertujuan untuk mengeksplorasi berbagai sumber penelitian sebelumnya yang relevan. Pemahaman terhadap kajian teori dilakukan untuk memahami konsep-konsep yang berkaitan dengan bidang penelitian yang dipilih. Luaran dari tahap ini adalah tinjauan pustaka.

3.2.2 Perancangan permasalahan penelitian

Perancangan permasalahan penelitian dilakukan dengan menelaah secara sistematis berbagai studi terdahulu untuk menemukan aspek-aspek yang masih belum terselesaikan atau belum diteliti secara optimal. Melalui proses ini, dilakukan penilaian terhadap kelemahan metode, keterbatasan dataset, ketidaktepatan teknik evaluasi, serta peluang pengembangan pendekatan baru yang belum dimanfaatkan dalam penelitian sebelumnya. Hasil analisis tersebut kemudian digunakan untuk merumuskan permasalahan inti yang perlu dijawab dalam penelitian ini.

3.2.3 Data Collecting

Tahap pengumpulan data bertujuan untuk memperoleh dataset yang digunakan sebagai dasar pengembangan model deteksi malware berbasis GRU. Pada penelitian ini, dataset diperoleh dari platform Kaggle, yang menyediakan dataset keamanan jaringan IoT yang dipublikasikan secara terbuka untuk keperluan penelitian.

Dataset yang digunakan terdiri dari CIC-IoT2023, N-BaIoT, dan Edge-IIoTset. Ketiga dataset tersebut dipilih karena merepresentasikan berbagai pola

serangan pada lingkungan IoT serta memiliki karakteristik fitur berdimensi tinggi (high-dimensional features), sehingga relevan untuk menguji efektivitas teknik reduksi dimensi yang menjadi fokus penelitian ini. Selain itu, penggunaan lebih dari satu dataset bertujuan untuk menguji generalisasi metode yang diusulkan, sehingga performa model tidak hanya optimal pada satu dataset tertentu, tetapi konsisten pada variasi karakteristik trafik jaringan IoT yang berbeda.

Dataset diunduh dalam format asli tanpa modifikasi struktur pada tahap pengumpulan untuk menjaga integritas data. Setelah diperoleh, dilakukan identifikasi awal terhadap struktur fitur, jumlah kelas, serta distribusi label guna memastikan kesesuaian dataset dengan kebutuhan penelitian dan sebagai dasar perancangan tahap preprocessing serta eksperimen pemodelan. Deskripsi rinci masing-masing dataset disajikan pada Bab IV.

3.2.4 Data Preprocessing

Data preprocessing dilakukan untuk memastikan data yang digunakan dalam penelitian ini berada dalam kondisi bersih dan siap digunakan dalam proses pelatihan model. Tahapan ini penting untuk meminimalkan gangguan numerik, mengurangi potensi bias, serta menjaga kualitas fitur sebelum dilakukan reduksi dimensi dan pemodelan menggunakan GRU.

Pada tahap awal, atribut non-numerik yang tidak relevan terhadap proses klasifikasi dihapus, karena model GRU dalam penelitian ini menggunakan masukan dalam bentuk numerik. Kolom label tetap dipertahankan sebagai penanda kelas. Selanjutnya dilakukan pemeriksaan terhadap *missing values*. Baris data yang

memiliki proporsi nilai kosong lebih dari 50% dihapus, dengan pertimbangan bahwa kehilangan informasi dalam jumlah besar dapat memengaruhi representasi fitur secara signifikan. Pendekatan ini dipilih agar data yang digunakan tetap memiliki kualitas informasi yang memadai tanpa menambahkan estimasi nilai yang berpotensi menimbulkan bias.

Nilai tak hingga (infinite values) yang ditemukan dalam dataset terlebih dahulu dikonversi menjadi NaN, kemudian dihapus untuk menghindari gangguan pada proses komputasi dan normalisasi. Selain itu, dilakukan pula pengecekan dan penghapusan data duplikat untuk memastikan tidak terjadi pengulangan data yang dapat memengaruhi proses pembelajaran model.

Setelah proses pembersihan selesai, dilakukan normalisasi menggunakan MinMaxScaler dengan rentang nilai 0 hingga 1. Normalisasi ini bertujuan untuk menyamakan skala antar fitur sehingga proses pelatihan model menjadi lebih stabil. Untuk menjaga validitas eksperimen dan menghindari data leakage, proses fitting scaler dilakukan hanya pada data latih, kemudian hasil transformasinya diterapkan pada data validasi dan data uji. Seluruh tahapan preprocessing tersebut dilakukan secara sistematis sebagai dasar sebelum memasuki tahap reduksi dimensi dan pengembangan model.

3.2.5 Data Splitting

Dataset dibagi menggunakan metode stratified splitting agar proporsi kelas pada setiap subset tetap sesuai dengan distribusi data asli. Pembagian dilakukan ke dalam tiga bagian, yaitu 60% sebagai data latih *(training)*, 20% sebagai data

validasi *(validation),* dan 20% sebagai data uji *(testing).* Skema 60:20:20 digunakan untuk memberikan porsi data yang cukup pada tahap pelatihan, sekaligus menyediakan data validasi untuk pemantauan performa model serta data uji untuk evaluasi akhir.

Karakteristik dataset trafik IoT umumnya tidak seimbang, di mana jumlah data normal jauh lebih banyak dibandingkan data serangan. Oleh karena itu, dilakukan penyeimbangan kelas pada data latih dan data validasi menggunakan Random UnderSampling (RUS). Teknik RUS dipilih dibandingkan metode oversampling seperti SMOTE karena penelitian ini berfokus pada evaluasi pengaruh teknik reduksi dimensi terhadap performa model. Penggunaan oversampling berpotensi menghasilkan data sintetis yang dapat memengaruhi distribusi fitur dan mengaburkan efek asli dari teknik reduksi dimensi yang diuji. Dengan menggunakan RUS, penelitian ini mempertahankan karakteristik asli data tanpa menambahkan sampel buatan, sehingga hasil eksperimen lebih merefleksikan dampak murni dari pendekatan reduksi fitur yang diterapkan.

Proses pembagian data dilakukan sebelum tahap pelatihan model untuk memastikan tidak terjadi tumpang tindih antar subset. Dengan demikian, setiap subset bersifat independen dan tidak menimbulkan kebocoran informasi (data leakage) antara data latih, validasi, dan uji. Setelah pembagian data selesai, dilakukan normalisasi fitur numerik menggunakan MinMaxScaler dengan rentang nilai 0 hingga 1. Proses fitting dilakukan hanya pada data latih, kemudian transformasi yang sama diterapkan pada data validasi dan data uji. Pendekatan ini

dilakukan untuk menjaga objektivitas pelatihan model dan memastikan bahwa informasi dari data uji tidak memengaruhi proses pembelajaran.

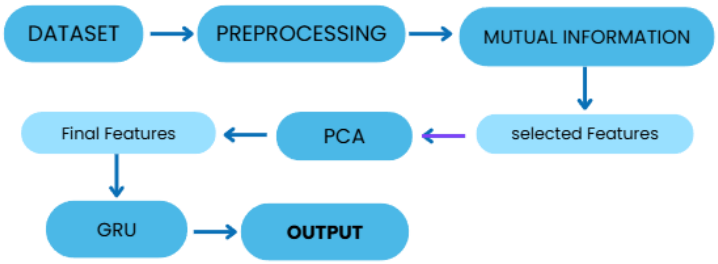
3.2.6 Modelling

Tahap pemodelan merupakan bagian inti dalam penelitian ini, karena pada tahap ini dilakukan pengujian terhadap pengaruh teknik reduksi dimensi terhadap kinerja model GRU dalam mendeteksi malware IoT. Proses pemodelan diawali dengan penentuan jumlah fitur optimal (K) untuk setiap metode reduksi dimensi yang digunakan, yaitu PCA, MI, serta pendekatan Hybrid yang menggabungkan MI dan PCA.

Penentuan nilai K dilakukan melalui proses tuning menggunakan data validasi. Untuk setiap kandidat jumlah fitur, model GRU dilatih menggunakan data latih, kemudian performanya dievaluasi pada data validasi menggunakan metrik F1-score. Metrik ini digunakan karena mampu memberikan keseimbangan antara precision dan recall, terutama pada data dengan distribusi kelas yang tidak seimbang. Nilai K yang menghasilkan F1-score terbaik pada data validasi kemudian dipilih sebagai konfigurasi akhir. Selain ketiga metode reduksi dimensi tersebut, model juga diuji dalam skenario tanpa reduksi fitur sebagai baseline pembanding. Dengan demikian, penelitian ini membandingkan empat skenario, yaitu baseline, PCA, MI, dan Hybrid.

Pendekatan Hybrid dalam penelitian ini merujuk pada kombinasi dua teknik reduksi dimensi yang dilakukan secara berurutan. Tahap pertama adalah seleksi fitur menggunakan Mutual Information untuk memilih fitur yang paling relevan terhadap label. Selanjutnya, fitur terpilih tersebut diproses kembali menggunakan

PCA untuk menghasilkan representasi fitur yang lebih ringkas sebelum dimasukkan ke dalam model GRU. Dengan demikian, arsitektur Hybrid yang dimaksud merupakan alur pemrosesan fitur (pipeline) sebelum tahap klasifikasi, bukan perubahan pada struktur layer jaringan. Alur pemodelan ini ditunjukkan pada Gambar 3.3.



Gambar 3.3 Alur Hybrid

Arsitektur GRU yang digunakan terdiri dari dua lapisan GRU berurutan dengan masing-masing 128 dan 64 unit. Untuk menjaga stabilitas proses pelatihan, digunakan mekanisme dropout, recurrent dropout, serta Batch Normalization. Setelah lapisan rekuren, ditambahkan satu lapisan Dense dengan aktivasi ReLU dan dropout, kemudian diakhiri dengan lapisan output beraktivasi sigmoid untuk klasifikasi biner.

Model dikompilasi menggunakan optimizer Adam dengan learning rate 0,001 dan loss function Binary Crossentropy. Proses pelatihan dilakukan dengan batas maksimum 20 epoch. Nilai tersebut digunakan sebagai batas atas pelatihan, sementara mekanisme Early Stopping diterapkan dengan memantau nilai validation loss dan patience sebesar 3 epoch. Dengan mekanisme ini, pelatihan akan berhenti secara otomatis apabila tidak terjadi peningkatan performa pada data validasi, sehingga membantu menjaga efisiensi pelatihan sekaligus meminimalkan risiko

overfitting. Seluruh eksperimen dijalankan menggunakan framework TensorFlow dengan pengaturan random seed 42 untuk menjaga konsistensi hasil antar percobaan.

3.2.7 Evaluasi

Evaluasi dilakukan untuk menilai kinerja model serta memastikan bahwa perbedaan performa antar skenario dapat dipertanggungjawabkan secara ilmiah. Pengujian dilakukan pada data uji (test set) yang tetap mempertahankan distribusi kelas aslinya. Pendekatan ini digunakan agar hasil evaluasi mencerminkan kondisi nyata pada lingkungan IoT, di mana trafik normal umumnya lebih dominan dibandingkan trafik serangan.

Kinerja model dievaluasi menggunakan beberapa metrik yang umum digunakan pada klasifikasi dengan data tidak seimbang, yaitu accuracy, precision, recall, F1-score, dan specificity. Accuracy digunakan untuk melihat performa keseluruhan model, sedangkan precision dan recall memberikan gambaran kemampuan model dalam mendeteksi kelas serangan secara lebih spesifik. F1-score digunakan sebagai metrik utama karena mempertimbangkan keseimbangan antara precision dan recall. Specificity turut dihitung untuk melihat kemampuan model dalam mengidentifikasi kelas normal secara benar. Hasil prediksi dari empat skenario pemodelan, yaitu baseline, PCA, MI, dan Hybrid, dibandingkan secara komparatif untuk melihat pengaruh masing-masing teknik reduksi dimensi terhadap performa model GRU.

Untuk memastikan bahwa perbedaan performa yang diperoleh tidak terjadi secara kebetulan, dilakukan Uji McNemar pada pasangan model yang

dibandingkan. Uji McNemar dipilih karena dirancang untuk mengevaluasi perbedaan performa dua classifier yang diuji pada dataset yang sama dengan pendekatan data berpasangan. Berbeda dengan uji statistik berbasis rata-rata seperti t-test yang memerlukan asumsi distribusi tertentu, Uji McNemar secara khusus menganalisis perbedaan pada prediksi yang tidak konsisten antara dua model.

Dengan demikian, uji ini lebih sesuai untuk mengevaluasi apakah peningkatan performa antara baseline, PCA, MI, dan Hybrid benar-benar signifikan secara statistik pada level prediksi individu, bukan hanya pada nilai metrik agregat. Pengujian dilakukan dengan membentuk tabel kontingensi 2×2 berdasarkan hasil prediksi kedua model yang dibandingkan. Nilai p yang diperoleh kemudian diinterpretasikan dengan tingkat signifikansi 0,05. Apabila nilai $p < 0,05$, maka perbedaan performa antara kedua model dianggap signifikan secara statistik.