

## **BAB III**

### **METODOLOGI PENELITIAN**

#### **3.1 Jenis Penelitian**

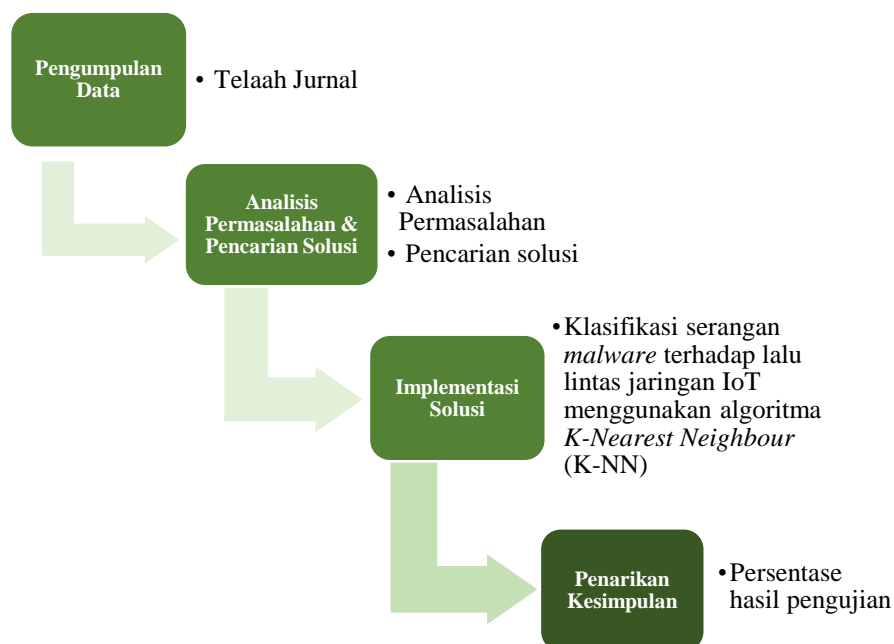
Jenis dari penelitian ini adalah penelitian eksperimental dengan pendekatan penelitian secara kualitatif. Penelitian eksperimen (*Experimental Research*) adalah suatu penelitian yang berusaha mencari pengaruh variabel tertentu terhadap variabel lain dalam kondisi yang terkontrol secara ketat (Azqiera, 2018). Penelitian kualitatif adalah penelitian yang bermaksud untuk memahami fenomena tentang apa yang dialami oleh subjek penelitian misalnya perilaku, persepsi, motivasi, tindakan, dll secara holistic, dan dengan cara deskripsi dalam bentuk kata-kata dan bahasa, pada suatu konteks khusus yang alamiah dan dengan memanfaatkan berbagai metode alamiah (Hidayat, 2012). Penelitian ini berjenis eksperimental karena objek yang diteliti diproses melalui suatu eksperimen yang menggunakan algoritma *K-Nearest Neighbour (K-NN)* sebagai *Classifier*.

#### **3.2 Objek dan Variabel Penelitian**

Objek yang diteliti pada penelitian ini adalah *dataset* dari lalu lintas jaringan IoT. Variabel yang digunakan pada penelitian ini adalah algoritma *K-Nearest Neighbour (K-NN)*.

### 3.3 Tahapan Penelitian

Gambar 3.1 merupakan tahap penelitian yang dimulai dari proses pengumpulan data, analisis permasalahan dan pencarian solusi, implementasi solusi, sampai proses penarikan kesimpulan:



Gambar 3.1 Tahapan Penelitian

#### 3.3.1 Pengumpulan Data

Data yang merupakan penunjang penelitian dapat diperoleh melalui studi literatur. Studi literatur berisi uraian tentang teori, temuan, dan bahan penelitian lainnya yang diperoleh dari jurnal nasional ataupun internasional yang berupa *survey paper* dan *technical paper*. Studi literatur yang digunakan untuk menunjang penelitian ini adalah jurnal terkait *data mining* dan algoritma *machine learning* untuk digunakan pada proses klasifikasi data, *dataset* yang layak digunakan

berjumlah 20 *dataset* dengan kolom atau atribut yang sama lalu *dataset* tersebut dilakukan data *preprocessing* untuk membersihkan *dataset* dari data yang kosong atau *missing values*, setelah *dataset* sudah melewati data *preprocessing* kemudian dilakukan *export dataset* menjadi sebuah file berekstensi .CSV dan akan dijadikan sebagai bahan penelitian.

### 3.3.2 Analisis Permasalahan dan Pencarian Solusi

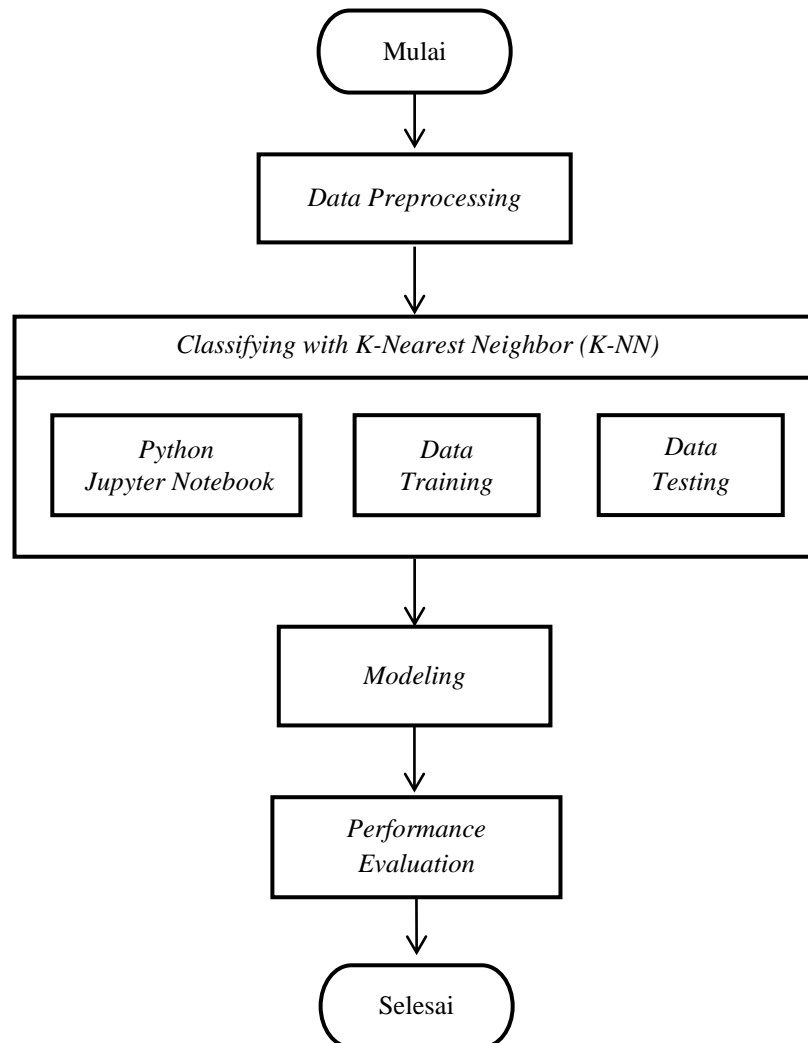
Tahapan analisis permasalahan dan pencarian solusi merupakan tahap pengembangan yang dilakukan pasca pengumpulan data. Masalah yang ditemukan pada proses studi literatur, kemudian dikaji dan dicarikan solusinya berdasarkan perkembangan ilmu pengetahuan dan teknologi yang ada. Rentannya keamanan dan banyaknya anomali pada lalu lintas jaringan IoT yang disebabkan serangan *malware* menjadi sebuah masalah yang harus dicari solusinya. Solusi yang dipilih adalah dilakukan pengklasifikasian serangan *malware* menggunakan algoritma *K-Nearest Neighbour (K-NN)* berdasarkan anomali *traffic* pada lalu lintas jaringan IoT.

Selain proses klasifikasi, pada penelitian ini juga dilakukan proses prediksi data lalu lintas jaringan IoT menggunakan model yang sudah dilakukan proses *training*, sehingga model dapat memprediksi *malware* menjadi 2 kategori yaitu *benign* dan *malicious*.

### 3.3.3 Implementasi Solusi

Tahap awal pada proses implementasi solusi adalah Data *preprocessing* dibagi menjadi 2 bagian, yaitu: data *cleaning* dan data *reduction*. Data *cleaning* adalah proses pembersihan data *incomplete* pada *attribute* di *dataset* untuk membuat data menjadi lebih konsisten. Sedangkan, data *reduction* adalah proses untuk menghapus data pada *attribute* yang kurang dominan sehingga data bisa dikurangi, namun tetap menghasilkan data yang akurat. Tahap kedua adalah *Classifying with K-Nearest Neighbor (K-NN)* yang merupakan proses klasifikasi serangan *malware* pada data lalu lintas jaringan IoT dikategorikan bersifat *benign* atau *malicious*. Tahap terakhir adalah tahap *Performance Evaluation*, setelah pembuatan model maka langkah selanjutnya adalah melakukan evaluasi dengan *performance evaluation*. *Performance evaluation* berguna untuk menguji performa dari *classifier*, *recall*, *precision*, dan *accuracy*. *Recall* adalah kumpulan data positif yang diklasifikasikan dengan benar sebagai data positif. *Precision* adalah kumpulan data yang diklasifikasikan sebagai positif yang benar – benar positif, *accuracy* adalah ketepatan klasifikasi data.

Gambar 3.2 merupakan *flowchart* yang menggambarkan secara utuh tahap implementasi solusi.



Gambar 3.2 *Flowchart* tahap implementasi solusi

### 3.3.4 Penarikan Kesimpulan

Penarikan kesimpulan merupakan tahap akhir dari penelitian ini dimana model yang sudah melewati proses *training* model menggunakan algoritma *K-Nearest Neighbour* (K-NN) yang dapat memprediksi data-data anomali *traffic* kedalam 2 kategori seperti *benign* dan *malicious*. Model tersebut di *training*

menggunakan 20 dataset anomali *traffic* yang sudah dipilih, data tersebut dibagi menjadi 60% untuk data *training* dan 40% untuk data *testing* yang menghasilkan nilai akurasi 94% sehingga dapat digunakan untuk memprediksi data-data anomali *traffic* selanjutnya.