



**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

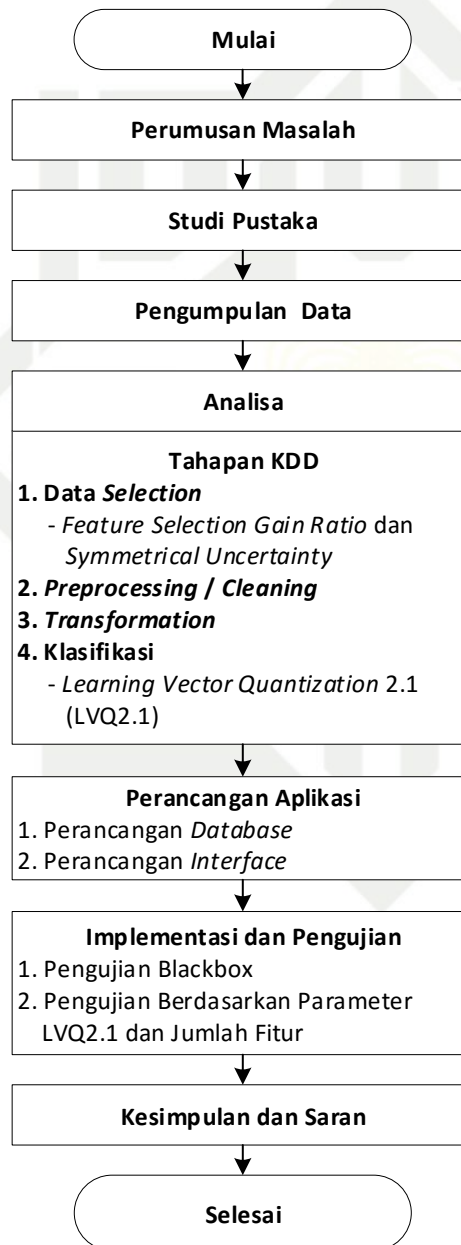
- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## BAB III

### METODOLOGI PENELITIAN

Metodologi penelitian merupakan tahapan sistematis yang akan dilaksanakan pada sebuah penelitian untuk menyelesaikan permasalahan pada penelitian tersebut. Adapun tahapan yang dilakukan pada penelitian ini dapat dilihat pada Gambar 3.1 dibawah ini:



**Gambar 3.1 Tahapan Metodologi Penelitian**



### 3.1 Perumusan Masalah

Penelitian ini merumuskan sebuah permasalahan yaitu bagaimana mengukur tingkat akurasi dari penerapan *feature selection* menggunakan *Gain Ratio* dan *Symmetrical Uncertainty* serta diklasifikasikan dengan metode *Learning Vector Quantization 2.1* (LVQ2.1) pada dataset NSL KDD.

### 3.2 Studi Pustaka

Tahapan ini dilakukannya pengumpulan informasi melalui penelitian terdahulu, buku-buku serta jurnal yang berkaitan dengan penelitian ini seperti, dataset NSL-KDD, *feature selection*, dan klasifikasi menggunakan *Learning Vector Quantization 2.1* (LVQ2.1).

### 3.3 Pengumpulan Data

Data yang menjadi penelitian ini adalah dataset NSL-KDD yang dapat diunduh pada situs [https://github.com/defcom17/NSL\\_KDD](https://github.com/defcom17/NSL_KDD). Dataset ini memiliki 41 parameter (atribut) dan memiliki 5 jenis kelas yang terdiri dari 1 jenis kelas normal dan 4 jenis serangan. Pada penelitian ini jenis kelas normal tidak digunakan karena tidak termasuk kedalam kategori serangan.

### 3.4 Analisa

Setelah dilakukan pengumpulan data maka tahapan selanjutnya adalah dilakukannya analisa terhadap data tersebut dengan *Knowledge Discovery in Database* (KDD). Tahapan KDD adalah sebagai berikut :

#### 3.4.1 Data Selection

Tahapan ini dilakukannya proses pemilihan (seleksi) data yang berperan penting dalam menentukan jenis serangan menggunakan metode *feature selection* (FS). Pada penelitian ini metode FS yang digunakan adalah kombinasi *gain ratio* dan *symmetrical uncertainty*.

Hasil dari tahapan ini yaitu diperolehnya nilai bobot pada masing-masing atribut/fitur dan kemudian dilakukan perangkingan berdasarkan nilai tertinggi ke terendah. Dalam menentukan atribut yang akan digunakan untuk klasifikasi, pada penelitian ini ditetapkan *threshold* yang ditentukan berdasarkan nilai bobot hasil



perangkingan pada FS *gain ratio* dan *symmetrical uncertainty*. Kemudian mengkombinasikan kedua FS tersebut untuk memperoleh atribut yang akan digunakan untuk klasifikasi.

### 3.4.2 Preprocessing / Cleaning

Tahapan ini dilakukannya proses pembersihan/penghapusan terhadap data, seperti data *missing value*, duplikasi data, data yang inkonsisten. Pada dataset NSL-KDD ini tidak terdapat data yang bersifat *missing value*, duplikasi data, maupun data yang inkonsisten.

### 3.4.3 Transformation

Tahapan ini dilakukannya proses merubah nilai atribut yang berbentuk simbol atau huruf menjadi angka. Pada dataset NSL-KDD yang digunakan pada penelitian ini, dari 41 atribut terdapat 3 atribut yang bernilai huruf dan akan ditransformasi, yaitu *protocol\_type*, *service*, *flag*. Kemudian merubah nilai-nilai angka tersebut menjadi berkisar antara 0-1 (normalisasi) menggunakan persamaan (2.14).

### 3.4.4 Klasifikasi

Tahapan ini diterapkannya metode klasifikasi untuk mengolah data yang ada menggunakan *Learning Vector Quantization 2.1* (LVQ2.1). Langkah-langkah yang dilakukan untuk klasifikasi metode LVQ2.1 diantaranya adalah (Budianita, 2013) :

1. X, vektor-vektor pelatihan ( $X_1, \dots, X_i, \dots, X_n$ ).
2. T, kategori atau kelas yang benar untuk vektor-vektor pelatihan.
3.  $W_j$ , vektor bobot pada unit keluaran ke-j ( $W_{1j}, \dots, W_{ij}, \dots, W_{nj}$ ).
4.  $C_j$ , kategori atau kelas yang merepresentasikan oleh unit keluaran ke-j.
5. *Learning rate* ( $\alpha$ ),  $\alpha$  didefinisikan sebagai tingkat pembelajaran. Jika  $\alpha$  terlalu besar, maka algoritma akan menjadi tidak stabil, sebaliknya jika  $\alpha$  terlalu kecil, maka prosesnya akan terlalu lama. Nilai  $\alpha$  adalah  $0 < \alpha < 1$ .
6. Nilai pengurangan *learning rate* ( $\alpha$ ), yaitu penurunan tingkat pembelajaran. Pengurangan nilai  $\alpha$  yang digunakan pada penelitian ini adalah sebesar 0.1 dengan rumus  $= \alpha - (0.1 * \alpha)$ .



#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

7. Nilai minimal *learning rate* (Min  $\alpha$ ), yaitu minimal nilai tingkat pembelajaran yang masih diperbolehkan. Nilai Min  $\alpha$  yang digunakan pada penelitian adalah sebesar 0.01.
8. Nilai *window* ( $\epsilon$ ), yaitu nilai yang digunakan sebagai daerah yang harus dipenuhi untuk memperbaharui vektor referensi pemenang dan *runner up* jika berada dikelas yang berbeda.
9. Jarak vektor referensi terdekat pertama dengan vektor  $x$  adalah  $d1$  dan jarak vektor referensi terdekat kedua dengan vektor  $x$  adalah  $d2$ .
10. Jika memenuhi kondisi *window* ( $\epsilon$ ), yakni pada Persamaan (2.10) maka vektor referensi yang masuk kedalam kelas tidak sama dengan vektor  $x$  akan diperbaharui menggunakan Persamaan (2.11).
11. Sedangkan vektor referensi yang masuk kedalam kelas yang sama dengan vektor  $x$  akan diperbaharui menggunakan Persamaan (2.12).

### 3.5 Perancangan Aplikasi

Perancangan aplikasi merupakan tahapan dilakukannya perancangan terhadap aplikasi yang akan dibuat, seperti perancangan *database* dan *interface*.

#### 3.5.1 Perancangan Database

Perancangan *database* dilakukan dengan tujuan untuk penyimpanan data yang dibuat dengan bentuk konseptual model.

#### 3.5.2 Perancangan Interface (Antarmuka)

Perancangan *interface* dilakukan dengan tujuan untuk memudahkan komunikasi antara pengguna dan aplikasi.

### 3.6 Implementasi dan Pengujian

Implementasi merupakan hasil dari tahap analisa sedangkan pengujian yaitu uji coba tingkat keakuratan kasus dengan metode yang telah diterapkan.

#### 3.6.1 Implementasi

Implementasi aplikasi merupakan tahapan yang menjelaskan tentang penerapan aplikasi yang telah dianalisa dan dirancang sebelumnya. Berikut adalah spesifikasi implementasi aplikasi yang digunakan :





#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

#### 1. *Hardware:*

CPU	: Intel® Core™ i3-3120M
Memori (RAM)	: 4 GB
HDD	: 500GB

#### 2. *Software:*

<i>Operating System</i>	: Windows 10
Bahasa Pemrograman	: PHP
DBMS	: MySQL
<i>Tools</i>	: Notepad++
<i>Web Browser</i>	: Google Chrome

### 3.6.2 Pengujian

Pengujian pada penelitian ini meliputi :

1. Pengujian *Blackbox*
2. Pengujian akurasi berdasarkan parameter LVQ2.1
3. Pengujian akurasi berdasarkan jumlah fitur / atribut.

### 3.7 Kesimpulan dan Saran

Tahapan ini menjelaskan tentang kesimpulan yang dapat diambil dari hasil penelitian yang telah dilakukan, serta memberikan saran kepada pembaca untuk mengembangkan penelitian agar menjadi lebih baik.