

Hak Cipta Dilindungi Undang-Undang

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

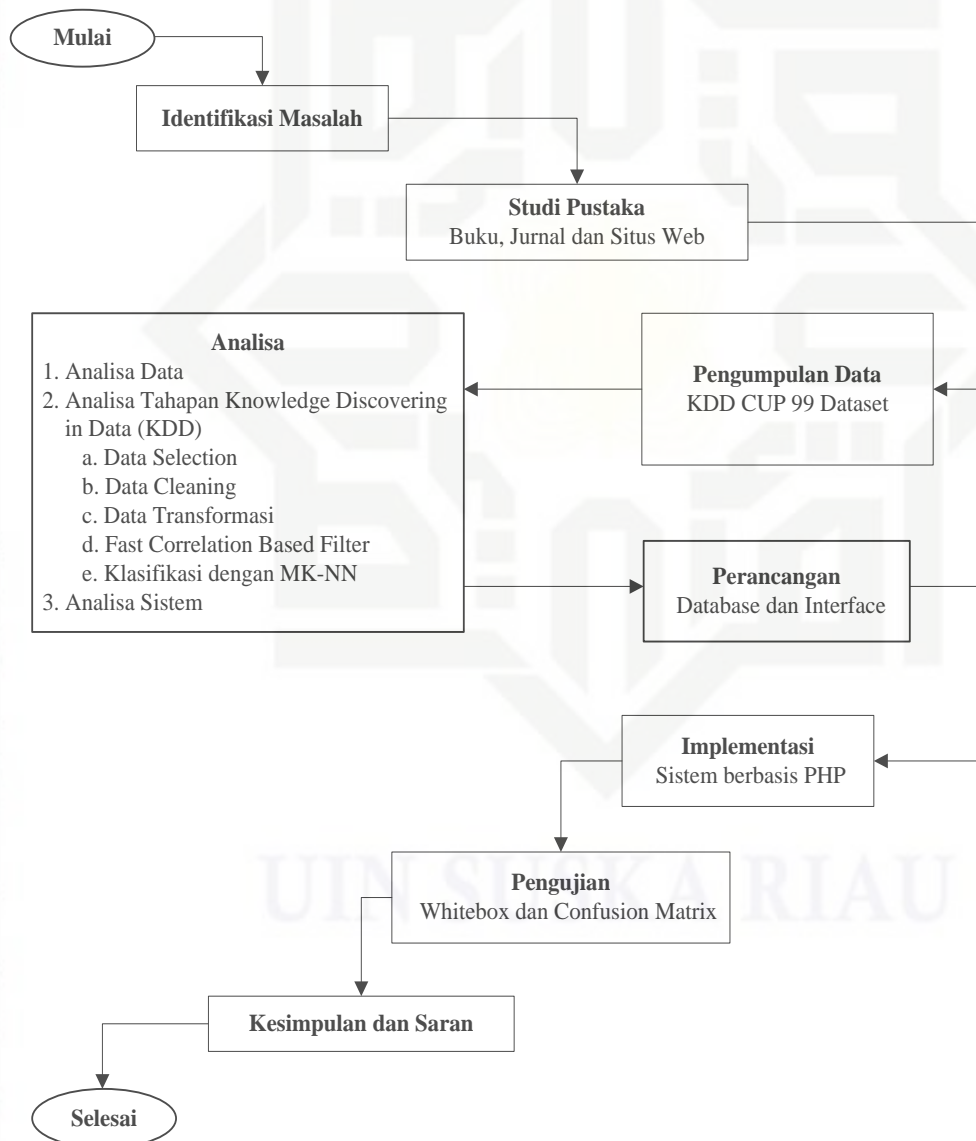
- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

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 adalah tahapan yang akan dilakukan dalam melakukan penelitian agar dapat memenuhi tujuan sesuai dengan yang diharapkan. Tahapan penelitian yang akan dilakukan :



Gambar 3.1 Tahapan Metodologi Penelitian

3.4.1 Analisa Data

Tahapan ini merupakan tahapan dimana data-data diperlukan dikumpulkan yang kemudian akan diolah dengan metode klasifikasi yang sebelumnya dilakukan praproses dengan seleksi fitur. Adapun penjelasan data yang akan digunakan adalah sebagai berikut :

1. Jumlah dataset 10% terdiri dari 494.020 *record*.
2. Mempunyai 41 fitur atau atribut.
3. Pengelompokan klasifikasi berjumlah 5 kelas.

3.4.2 Analisa Tahapan *Knowledge Discovering in Data* (KDD)

Tahapan ini menjelaskan proses pengolahan data pada dataset KDD CUP 99 dan berikut penjelasan dari tahapan yang dilalui :

a. *Data Selection*

Selanjutnya akan dilakukan proses *selection* terhadap data penelitian. Ada 5 kelas yang terbagi yaitu *Normal*, *DoS*, *U2R*, *R2L*, dan *Probe*.

b. *Data Cleaning*

Proses *cleaning* data atau pembersihan dilakukan terhadap data yang redudan atau ganda, inkonsisten, *missing value* dan *outlier* data. Hal tersebut perlu dilakukan agar tidak mempengaruhi performa proses klasifikasi yang dilakukan. Pembersihan data yang dilakukan juga menghapus fitur yang tidak diperlukan dalam proses klasifikasi nantinya.

c. *Data Transformation*

Tahap selanjutnya adalah transformasi data. Pada tahap ini data input yang bertipe text akan terlebih dahulu diubah kedalam bentuk numerik. Selanjutnya transformasi yang akan dilakukan adalah normalisasi. Tujuan dari normalisasi ini agar data berada pada *range* [0-1] sehingga sebaran datanya tidak terlalu jauh.

d. *Fast Correlation Based Filter* (FCBF)

Pada tahap ini Fitur akan diseleksi dengan menggunakan algoritma *Information Gain* (persamaan 2.3). Setelah mendapatkan nilai IG, hitung

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.

Symmetrical Uncertainty (persamaan 2.4). Fitur yang terpilih dan akan digunakan dalam proses klasifikasi jika $SU > \delta$ (*Threshold*) dan nilai *Threshold* yang digunakan 0.

e. **Klasifikasi *Modified K-Nearest Neighbor* (MK-NN)**

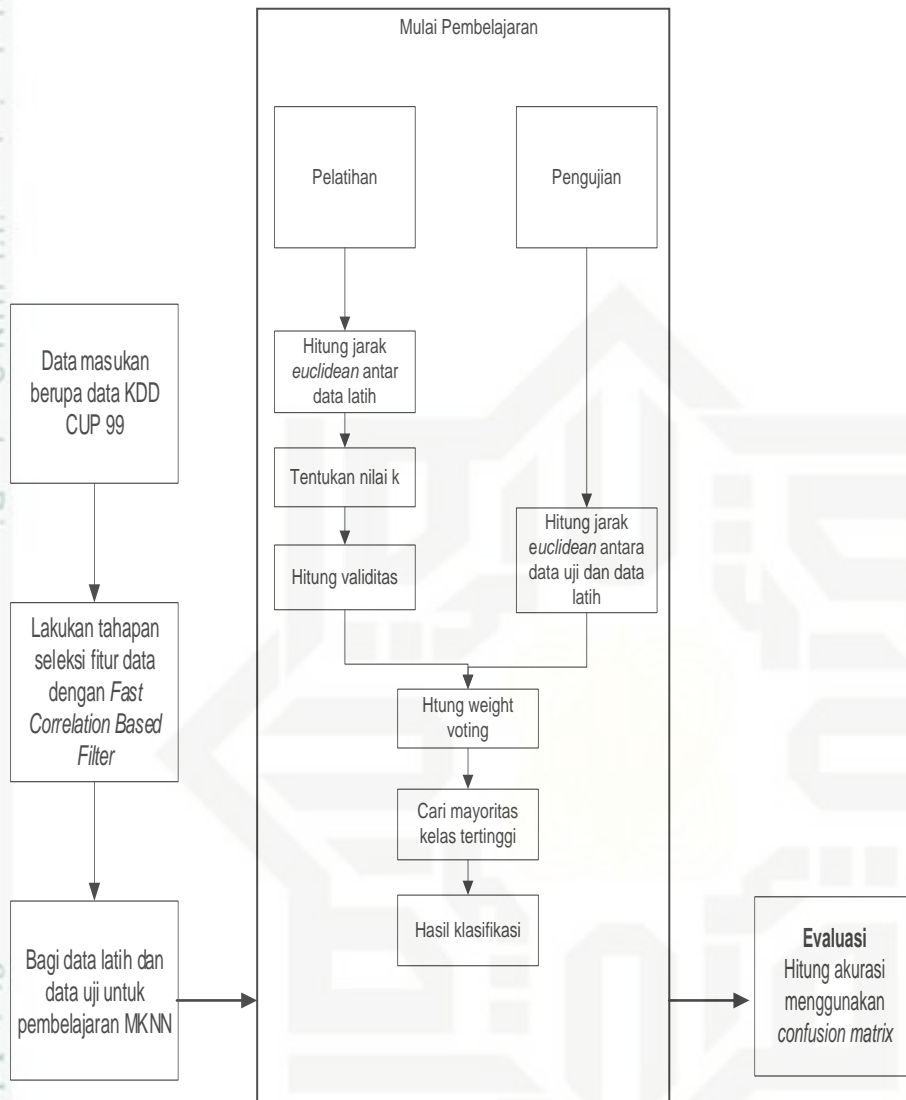
Tahapan ini adalah tahapan yang menerapkan metode dari *data mining* untuk mengolah data yang ada. Metode yang digunakan adalah metode klasifikasi *modified k-nearest Neighbor*. Metode ini melakukan klasifikasi serangan jaringan komputer pada dataset KDD CUP 99 berdasarkan data latih yang jaraknya paling dekat. Untuk mencari jarak tersebut dihitung menggunakan rumus *Euclidean* (Persamaan 2.5).

Pada tahapan ini metode *Modified k-Nearest Neighbor* juga menentukan nilai k , kemudian dihitung nilai validitas (persamaan 2.10) dari data latih dan *weight voting* (persamaan 2.12). Kemudian tentukan kelas mayoritas tertinggi dari nilai k yang ditentukan. Selanjutnya akan menghasilkan hasil klasifikasi untuk menentukan kelas pada data uji. *Output* yang didapat adalah klasifikasi hasil dari dataset *KDD CUP 99*.

3.4.3 **Analisa Sistem**

Analisa perancangan pada sistem klasifikasi serangan pada dataset KDD CUP 99 yang akan dibangun meliputi : *Context Diagram*, *Data Flow Diagram* (DFD), *Flowchart* dan *Entity Relationship Diagram* (ERD).

Adapun gambaran umum tahapan proses analisa sistem yang akan dilakukan pada penelitian ini dapat dijelaskan berdasarkan gambar berikut :



Gambar 3.2 Flowchart Metode *Fast Correlation Based Filter* dan *Modified K-Nearest Neighbor*

3.5 Perancangan

Setelah melakukan analisa terhadap data menggunakan metode MK-NN maka akan dilanjutkan ke tahap perancangan untuk pembangunan aplikasi. Berikut tahapan perancangan untuk pembangunan aplikasi :

1. Perancangan database

Rancangan *database* diperlukan untuk merancang data yang akan dikelola didalam sistem yang akan dibangun.

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.

2. Perancangan *Interface*

Untuk mempermudah komunikasi antara sistem dengan pengguna, maka perlu dirancang antarmuka (*interface*). Dalam perancangan *interface*, hal terpenting yang ditekankan adalah bagaimana menciptakan tampilan dan struktur menu yang baik dan mudah dimengerti oleh pengguna.

3.6 Implementasi

Implementasi dikembangkan berdasarkan hasil analisa dan perancangan yang telah dilakukan sebelumnya. Tahap ini menandakan bahwa sistem siap untuk dioperasikan pada keadaan yang sebenarnya, sehingga akan diketahui apakah hasil yang diperoleh dari sistem sesuai dengan tujuan yang diinginkan. Implementasi dari hasil analisa dan perancangan pada penelitian ini adalah sistem yang berbasis web. Perangkat pendukung yang dibutuhkan sebagai berikut:

1. Platform : Multiplatform
2. Bahasa Pemograman : PHP
3. DBMS : MySQL
4. Browser : Google Chrome, Mozilla Firefox.
5. Web Server : Apache
6. Sistem Operasi : Windows 10 64-bit
7. Processor : Intel Core i3 (2,53 GHz)
8. Memory : RAM 4 GB

3.7 Pengujian

Setelah dilakukan implementasi, maka dilakukan pengujian terhadap sistem yang telah dibuat. Dengan tujuan agar menjamin sistem yang dibuat sesuai dengan hasil analisa dan perancangan dan juga menemukan kesalahan-kesalahan yang mungkin terjadi. Pada pengujian ini ada beberapa hal yang dijadikan acuan, yaitu :

1. *Whitebox*

Whitebox testing adalah pengujian yang didasarkan pada pengecekan terhadap detail perancangan, menggunakan struktur kontrol dari desain program

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.

secara procedural untuk membagi pengujian ke dalam beberapa kasus pengujian. Secara sekilas dapat diambil kesimpulan *whitebox* testing merupakan petunjuk untuk mendapatkan program yang benar secara 100%. Pengujian dilakukan berdasarkan bagaimana suatu software menghasilkan output dari input . Pengujian ini dilakukan berdasarkan kode program.

2. **Blackbox**

Metode ini berpusat pada fungsional perangkat lunak yang digunakan.

Tujuan metode ini adalah menemukan kesalahan-kesalahan seperti :

- a. Fungsi-fungsi yang tidak sesuai, atau hilang.
- b. Kesalahan atau kekeliruan *interface*.
- c. Kesalahan performansi sistem.
- d. Kesalahan pengaksesan database atau struktur data yang digunakan.
- e. Kesalahan inisialisasi (proses mulai) atau terminasi (proses selesai/akhir).

3. **Confusion Matrix**

Confusion matrix adalah alat visualisasi yang digunakan pada supervised learning. Tiap kolom pada mariks adalah contoh kelas prediksi, sedangkan tiap baris mewakili kejadian di kelas yang sebenarnya. Confusion matrix berisi informasi *actual* dan prediksi pada sistem klasifikasi.

3.8 **Kesimpulan dan Saran**

Bagian kesimpulan merupakan tahap penentuan kesimpulan terhadap hasil pengujian yang telah dilakukan. Hal tersebut bertujuan untuk mengetahui apakah sistem yang dirancang dan dibangun telah memenuhi kriteria yang sesuai dengan kebutuhan serta dapat dioperasikan dengan baik sehingga dapat bermanfaat. Pada bagian saran berisi kemungkinan pengembangan yang dapat dilakukan terhadap penelitian tersebut.