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 IV
ANALISA DAN PERANCANGAN

Pada bab ini akan membahas tentang analisa dan perancangan proses seleksi fitur dengan menggunakan metode *Fast Correlation Based Filter* dan proses klasifikasi dengan menggunakan metode *Modified K-Nearest Neighbour*. Model klasifikasi dianalisa dan dirancang sebagai pondasi untuk membangun sistem klasifikasi berbasis *web*. Sementara tahap perancangan merupakan tahap kegiatan menentukan rincian sistem yang akan dibuat berdasarkan analisa pada tahap sebelumnya. Berikut pembahasannya:

4.1 Analisa Kebutuhan Data

Data yang digunakan pada penelitian ini adalah dataset KDD CUP 1999 yang diunduh dan dicatat secara manual melalui situs http://www.kdd.org/kdd-cup/view/kdd-cup-1999/Data dan http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html yang berjumlah 494.021 data (10% dataset). Selanjutnya data menjadi data inputan pada aplikasi. Atribut/fitur data yang digunakan dalam penelitian ini berjumlah 41 fitur yang dibagi kedalam tiga kelompok yaitu fitur dasar, fitur konten dan fitur trafik. Fitur basic (fitur nomor 1 sampai 9) merupakan hasil ekstraksi dari sistem *log tcpdump* dalam jaringan komputer. Fitur konten (fitur nomor 10 sampai 22) merupakan fitur-fitur yang diambil dari kegiatan yang berlangsung dalam sistem jaringan komputer. Sedangkan fitur trafik terbagi menjadi dua bagian, pertama terdiri dari fitur nomor 23 sampai 31 merupakan fitur trafik jaringan yang dihitung menggunakan waktu dua detik *time window*, dan kedua terdiri dari fitur nomor 32 sampai 41 dihitung menggunakan waktu dua detik *time window* dari tujuan ke host.

Tabel 4.1 Fitur dasar (*basic*) tiap-tiap koneksi TCP

No	Nama Fitur	Keterangan	Tipe Data
1	*duration*	Lama (detik) koneksi	Continuous
2	*protocol_type*	Tipe protokol (tcp, udp, dll)	discrete

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.

No	Nama Fitur	Keterangan	Tipe Data
3	*service*	*Network service* di *destination* (http, telnet, dll)	discrete
4	*flag*	Status, normal atau error dari koneksi	discrete
5	*scr_bytes*	Jumlah rata-rata byte, termasuk informasi header yang diterima oleh destination host	continuous
6	*dst_bytes*	Jumlah rata-rata byte termasuk informasi header yang diterima oleh source host	continuous
7	*land*	1 jika koneksi bersal dari/ke host yang sama/port; 0 jika tidak	discrete
8	*wrong_fragment*	Jumlah *fragment* yang salah	continuous
9	*urgent*	Jumlah paket yang *urgent*	continuous

Tabel 4.2 Fitur konten pada koneksi (*content based*) berdasarkan *knowledge domain*

No	Nama Fitur	Keterangan	Tipe Data
1	*hot*	Jumlah indicator "*hot*" secara beruntun	continuous
2	*num_failed_logins*	Jumlah dari percobaan login yang gagal	continuous
3	*logged_in*	1 jika berhasil login; 0 jika tidak	discrete
4	*num_compromised*	Jumlah kondisi "*compromised*"	continuous
5	*root_sheel*	1 jika *root shell* didapat, 0 sebaliknya	discrete
6	*su_attempted*	1 jika dilakukan percobaan perintah "*su root*", 0 sebaliknya	discrete
7	*num_root*	Jumlah "*root*" yang diakses	continuous
8	*num_file_creations*	Jumlah operasi pembuaan file	continuous
9	*num_shells*	Jumlah *prompts shell*	continuous

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.

No	Nama Fitur	Keterangan	Tipe Data
10	*num_access_files*	Jumlah operasi pada *access control files*	continuous
11	*num_outbound_cmds*	jumlah dari perintah outbound dalam sesi ftp	Continuous
12	*is_host_login*	1 jika login termasuk dalam daftar "*hot*" 0 sebaliknya	Discrete
13	*is_guest_login*	1 jika *login* adalah "*guest*", 0 sebaliknya	Discrete

Tabel 4.3 Fitur trafik dihitung dengan menggunakan jeda waktu dua detik (*time based traffic features* dan *host based traffic features*)

No	Nama Fitur	Keterangan	Tipe Data
1	*count*	Jumlah koneksi ke *host* yang sama dengan koneksi yang ada sekarang dalam rentang 2 detik	Continuous
2	*serror_rate*	% dari keneksi yang terdapat "SYN" *error*	Continuous
3	*rerror_rate*	% dari koneksi yang terdapat "REJ" *error*	Continuous
4	*same_srv_rate*	% dari koneksi ke *service* yang sama	Continuous
5	*diff_srv_rate*	% dari koneksi ke *service* yang berbeda	Continuous
6	*srv_count*	Jumlah koneksi ke *service* yang sama terakhir	Continuous
7	*srv_serror_rate*	% dari koneksi yang terdapat "SYN"	Continuous
8	*srv_rerror_rate*	*Error* % dari koneksi yang terdapat	Continuous
9	*srv_diff_host_rate*	% dari koneksi ke *host* yang berbeda	Continuous
10	*dst_host_count*	Jumlah koneksi ke *host* yang	Continuous

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.

No	Nama Fitur	Keterangan	Tipe Data
		sama dengan koneksi ke *host* yang sama sekarang dalam rentang 2 detik	
11	*dst_host_serror_rate*	% dari koneksi yang terdapat "SYN" *error*	continuous
12	*dst_host_rerror_rate*	% dari koneksi yang terdapat "REJ" *error*	continuous
13	*dst_host_same_srv_rate*	% dari koneksi ke *service* yang sama	continuous
14	*dst_host_diff_srv_rate*	% dari koneksi ke *service* yang berbeda	continuous
15	*dst_host_srv_count*	% dari koneksi yang terdapat "REJ"*error*	continuous
16	*dst_host_srv_serror_rate*	% dari koneksi yang terdapat "REJ"*error*	continuous
17	*dst_host_srv_rerror_rate*	% dari koneksi yang terdapat "REJ"*error*	continuous
18	*dst_host_srv_diff_host_rate*	% dari koneksi ke host yang berbeda	continuous
19	*dst_host_same_src_port_rate*	% dari koneksi ke *port service* yang sama	continuous

Pada total 494.020 data (tabel 2.3) yang diperoleh, terdapat data dengan kelas dos sebanyak 391.458 data, kelas r2l sebanyak 1.126 data, kelas u2r sebanyak 52 data, kelas probe sebanyak 4.107 data dan 97.277 data untuk kelas normal. Berdasarkan fitur yang telah dijelaskan pada tabel 4.1, 4.2 dan 4.3 di atas, maka contoh data yang digunakan pada penelitian ini akan dijabarkan pada tabel 4.4 di bawah ini (selengkapnya di lampiran A).

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.
© Hak cipta milik UIN Suska Riau State Islamic University of Sultan Syarif Kasim Riau

Tabel 4.4 Data Penelitian

No	duration	protocol type	service	flag	src bytes	dst bytes	land	....	kelas
1	0	tcp	http	sf	181	5450	0	....	normal
2	0	tcp	http	sf	239	486	0	....	normal
3	0	tcp	http	sf	235	1337	0	....	normal
4	0	tcp	http	sf	219	1337	0	....	normal
5	0	tcp	http	sf	217	2032	0	....	normal
6	0	tcp	http	sf	217	2032	0	....	normal
7	0	tcp	http	sf	212	1940	0	....	normal
8	0	tcp	http	sf	159	4087	0	....	normal
9	0	tcp	http	sf	210	151	0	....	normal
10	0	tcp	http	sf	212	786	0	....	normal
11	0	tcp	http	sf	210	624	0	....	normal
12	0	tcp	http	sf	177	1985	0	....	normal
13	0	tcp	http	sf	222	773	0	....	normal
14	0	tcp	http	sf	256	1169	0	....	normal
15	0	tcp	http	sf	241	259	0	....	normal
16	0	tcp	http	sf	260	1837	0	....	normal
17	0	tcp	http	sf	241	261	0	....	normal
18	0	tcp	http	sf	257	818	0	....	normal
19	0	tcp	http	sf	233	255	0	....	normal
20	0	tcp	http	sf	233	504	0	....	normal
....	....	....	....	....	....	....	....	....	....
494.021	0	tcp	private	rstr	0	0	0	....	u2r

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.

4.2 Tahapan Knowledge Discovering in Data

Pada tahapan ini dilakukan langkah-langkah proses data mining yang dimulai dari *selection, preprocessing, transformation, fast correlation based filter* dan klasifikasi menggunakan algoritma *modified k-nearest neighbor*. Berikut adalah tahapan-tahapan yang dilakukan :

4.2.1 Seleksi Data (*Data Selection*)

Tahap data *selection* merupakan pemilihan (seleksi) data yang akan digunakan dalam penelitian. Tahapan ini perlu dilakukan sebelum dilakukan tahap perhitungan. Seleksi yang dilakukan adalah dengan menghapus fitur-fitur yang tidak diperlukan untuk proses *mining.* Menurut Essra dkk, (2016) tidak semua fitur yang ada didalam dataset KDD CUP 99 memberikan kontribusi pada karakteristik trafik jaringan. Kayacik dkk, (2005) juga menyimpulkan bahwa tidak semua 41 fitur dibutuhkan untuk mengklasifikasikan jenis serangan. Fitur yang tidak digunakan untuk proses selanjutnya adalah *sevice* dan *flag*. Sehingga total fitur setelah seleksi data yang digunakan menjadi 39 atribut. Berdasarkan tabel 4.5, berikut ini adalah hasil proses *selection* yang telah dilakukan. (Selengkapnya di Lampiran B)

Tabel 4.5 Seleksi Data (*Data Selection*)

No	duration	src_bytes	dst_bytes	wrong_fragment	urgent	count	serror_rate	....	Kelas
1	0	181	5450	0	0	0	0	....	normal
2	0	239	486	0	0	0	0	....	normal
3	0	235	1337	0	0	0	0	....	normal
4	0	219	1337	0	0	0	0	....	normal
5	0	217	2032	0	0	0	0	....	normal
6	0	217	2032	0	0	0	0	....	normal
7	0	212	1940	0	0	0	0	....	normal
8	0	159	4087	0	0	0	0	....	normal

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.

No	duration	src_bytes	dst_bytes	wrong_fragment	urgent	count	serror_rate	....	Kelas
9	0	210	151	0	0	0	0	....	normal
10	0	212	786	0	0	1	0	....	normal
11	0	210	624	0	0	0	0	....	normal
12	0	177	1985	0	0	0	0	....	normal
13	0	222	773	0	0	0	0	....	normal
14	0	256	1169	0	0	0	0	....	normal
15	0	241	259	0	0	0	0	....	normal
16	0	260	1837	0	0	0	0	....	normal
17	0	241	261	0	0	0	0	....	normal
18	0	257	818	0	0	0	0	....	normal
19	0	233	255	0	0	0	0	....	normal
20	0	233	504	0	0	0	0	....	normal
21	0	256	1273	0	0	0	0	....	normal
22	0	234	255	0	0	0	0	....	normal
23	0	241	259	0	0	0	0	....	normal
24	0	239	968	0	0	0	0	....	normal
25	0	245	1919	0	0	0	0	....	normal
26	0	248	2129	0	0	0	0	....	normal
27	0	354	1752	0	0	0	0	....	normal
28	0	193	3991	0	0	0	0	....	normal
29	0	214	14959	0	0	0	0	....	normal
30	0	212	1309	0	0	0	0	....	normal
....	....	....	....	....	....	....	....	....	....
260	0	4	0	0	0	0	0	....	u2r

4.2.2 Praproses Data *(Data Preprocessing)*

Tahap data *processing* ini menggunakan proses *cleaning* data. Proses pembersihan terhadap data yang tidak konsisten, *missing value* atau data yang

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.

hilang, *outlier* atau data yang memiliki nilai yang berlebihan diantara data-data yang ada, serta data yang redudan atau ganda. Pada penelitian ini dilakukan proses *cleaning* dengan cara menghapus data yang redudan. Pada tahapan *cleaning* ini dilakukan melalui sistem. Caranya adalah dengan menyeleksi semua fitur yang ada dalam satu data kemudian dibandingkan dengan semua data yang ada. Apabila terdapat data yang persis sama, maka semua data yang ada tersebut dihapus dan hanya menyisakan satu *record* saja. Setelah dilakukan *cleaning*, data yang semula berjumlah 494.021 data menjadi berjumlah 145.585 data.

4.2.3 Transformasi Data (*Data Transformation*)

Setelah melakukan tahap preprocessing data, data yang akan digunakan pada transformasi sebanyak 260 data. Berdasarkan dengan keadaan kelas data yang seimbang (*Balance Class*) artinya tidak ada kelas yang lebih mendominasi. Sebanyak 52 data yang diambil, kemudian dikalikan dengan 5 fitur yang berbeda, sehingga masing–masing dari ke 5 fitur mendapatkan 52 jumlah data yang seimbang untuk masing-masing kelas. Sehingga diharapkan akan mampu menambah kompleksitas serta keakurasian dari model klasifikasi sistem yang akan dibangun dan menghindari selisih hasil rentang target yang berat sebelah karena jumlah data yang berbeda dan bervariasi. Pada tahap ini dilakukan transformasi data dengan menormalisasi (persamaan 2.1). Normalisasi dilakukan pada data yang memiliki sebaran yang jauh. Tujuannya adalah agar sebaran data berada pada rentang nilai [0-1]. Pada penelitian ini data yang dinormalisasikan adalah data dengan fitur *duration*, *protocol_type*, *scr_bytes*, *dst_bytes*, *wrong_fragment*, *urgent*, *count*, *serror_rate*, *same_srv_rate*, *dst_host_diff_ srv_rate*, *dst_host_srv_count*, *num_compromised* dan *root_sheel*. Sedangkan fitur lainnya tidak perlu dinormalisasikan karena rentang nilainya sudah berada pada [0-1]. Berdasarkan pada tabel 4.3, Berikut adalah proses perhitungan normalisasi:

Pada kolom *duration*, dari 260 data nilai minimumnya adalah 0 dan nilai maksimumnya adalah 718

v^i (data 1, kolom *duration*) = $v^i = \frac{v - min_a}{max_a - min_a} (new\_max_a - new\_min_a) + new\_min_a$

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.
© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau
UIN SUSKA RIAU

$$v^i = \frac{0-0}{(718-0)}(1-0)+0$$

$$v^i = \frac{0}{718}$$

$$v^i = \mathbf{0}$$

Pada kolom *scr_bytes*, dari 260 data, nilai minimumnya adalah 0 dan nilai maksimumnya adalah 54540

v^i (data 1, kolom *scr_bytes*) = $v^i = \frac{v-\ min_a}{max_a-\ min_a}\ (new\_max_a - new\_min_a) + new\_min_a$

$$v^i = \frac{181-0}{(54540-0)}(1-0)+0$$

$$v^i = \frac{181}{54540}$$

$$v^i = \mathbf{0.003}$$

Pada kolom *dst_bytes*, dari 260 data, nilai minimumnya adalah 0 dan nilai maksimumnya adalah 5155468

v^i (data 1, kolom *dst_bytes*) = $v^i = \frac{v-\ min_a}{max_a-\ min_a}\ (new\_max_a - new\_min_a) + new\_min_a$

$$v^i = \frac{5450-0}{(5155468-0)}(1-0)+0$$

$$v^i = \frac{5450}{5155468}$$

$$v^i = \mathbf{0.001}$$

Pada kolom *wrong_fragment*, dari 260 data, nilai minimumnya adalah 0 dan nilai maksimumnya adalah 3

v^i (data 1, kolom *wrong_fragment*) = $v^i = \frac{v-\ min_a}{max_a-\ min_a}\ (new\_max_a - new\_min_a) + new\_min_a$

$$v^i = \frac{0-0}{(3-0)}(1-0)+0$$

$$v^i = \frac{0}{3}$$

$$v^i = \mathbf{0}$$

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.

Perhitungan dilakukan pada seluruh *record* yang ada. Berdasarkan perhitungan di atas, hasil transformasi data dapat dilihat pada tabel 4.6 dibawah ini (selengkapnya di lampiran C)

Tabel 4.6 Data Transformation

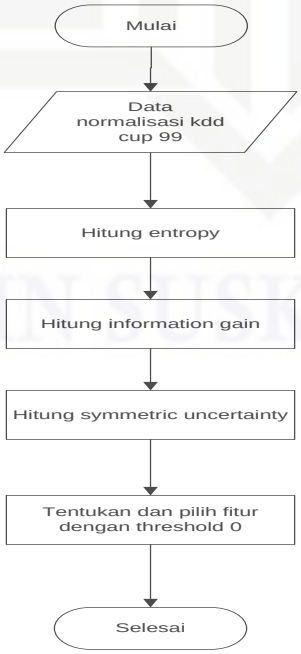
No	dura tion	src_byt es	dst_byt es	wrong _frag ment	urgent	count	serror _rate	....	jenis
1	0	0,0033	0,0011	0	0	0	0	....	normal
2	0	0,0044	0,0001	0	0	0	0	....	normal
3	0	0,0043	0,0003	0	0	0	0	....	normal
4	0	0,004	0,0003	0	0	0	0	....	normal
5	0	0,004	0,0004	0	0	0	0	....	normal
6	0	0,004	0,0004	0	0	0	0	....	normal
7	0	0,0039	0,0004	0	0	0	0	....	normal
8	0	0,0029	0,0008	0	0	0	0	....	normal
9	0	0,0039	0	0	0	0	0	....	normal
10	0	0,0039	0,0002	0	0	0,0556	0	....	normal
11	0	0,0039	0,0001	0	0	0	0	....	normal
12	0	0,0032	0,0004	0	0	0	0	....	normal
13	0	0,0041	0,0001	0	0	0	0	....	normal
14	0	0,0047	0,0002	0	0	0	0	....	normal
15	0	0,0044	0,0001	0	0	0	0	....	normal
16	0	0,0048	0,0004	0	0	0	0	....	normal
17	0	0,0044	0,0001	0	0	0	0	....	normal
18	0	0,0047	0,0002	0	0	0	0	....	normal
19	0	0,0043	0	0	0	0	0	....	normal
20	0	0,0043	0,0001	0	0	0	0	....	normal
21	0	0,0047	0,0002	0	0	0	0	....	normal
22	0	0,0043	0	0	0	0	0	....	normal
23	0	0,0044	0,0001	0	0	0	0	....	normal

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.

No	dura tion	src_byt es	dst_byt es	wrong _frag ment	urgent	count	serror _rate	....	jenis
24	0	0,0044	0,0002	0	0	0	0	....	normal
25	0	0,0045	0,0004	0	0	0	0	....	normal
26	0	0,0045	0,0004	0	0	0	0	....	normal
27	0	0,0065	0,0003	0	0	0	0	....	normal
28	0	0,0035	0,0008	0	0	0	0	....	normal
29	0	0,0039	0,0029	0	0	0	0	....	normal
30	0	0,0039	0,0003	0	0	0	0	....	normal
....	....	....	....	....	....	....	....	....	....
260	0	4	0	0	0	0	0		u2r

4.2.4 Seleksi Fitur dengan *Fast Correlation Based Filter* (FCBF)

Setelah melalui tahapan sebelumnya, pada bagian ini dijelaskan bagaimana fitur yang akan diseleksi dan digunakan nantinya sebagai proses klasifikasi. Pemilihan fitur menggunakan metode FCBF, algoritma ini akan dijelaskan pada *flowchart* yang diperlihatkan pada gambar 4.1 dibawah ini :



Gambar 4.1 Flowchart Seleksi Fitur Fast Correlation Based Filter

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.

Berikut keterangan dari *flowchart* seleksi fitur KDD CUP 99 dengan menggunakan metode *Fast Correlation Based Filter* :

1. Data Normalisasi KDD CUP 99

Data normalisasi KDD CUP 99 merupakan semua data KDD CUP 99 yang sudah dinormalisasi dan digunakan dalam penelitian. Data tersebut telah melewati proses *selection, preprocessing* dan *transformation* yang terdapat pada Tabel 4.6.

2. Menghitung Nilai Entropy

Proses perhitungan nilai *entropy* digunakan sebagai suatu parameter untuk mengukur heterogenitas (keberagaman) dari suatu kumpulan sampel data. Perhitungan *entropy* menggunakan persamaan 2.1, berikut langkah – langkah perhitungan *entropy* :

$$Entropy(S) = \sum_{i}^{c} - \ p_i \ log_2 \ p_i$$

Entropy(Total) $= (-\frac{52}{260} * \log_2(\frac{52}{260})) + (-\frac{208}{260} * \log_2(\frac{208}{260})) = 0.7219$

Entropy(Total) = 0.7219

Entropy duration

Entropy (Total , 0) $= (-\frac{137}{189} * \log_2(\frac{137}{189})) + (-\frac{52}{189} * \log_2(\frac{52}{189})) = 0.8487$

Entropy (Total , 1) $= (-\frac{1}{1} * \log_2(\frac{1}{1})) + (-\frac{0}{1} * \log_2(\frac{0}{1})) = 0$

Entropy src_bytes

Entropy (Total , 0) $= (-\frac{90}{90} * \log_2(\frac{90}{90})) + (-\frac{0}{90} * \log_2(\frac{0}{90})) = 0$

Entropy (Total , 1) $= (-\frac{9}{9} * \log_2(\frac{9}{9})) + (-\frac{0}{9} * \log_2(\frac{0}{9})) = 0$

Entropy dst_bytes

Entropy (Total , 0) $= (-\frac{127}{130} * \log_2(\frac{127}{130})) + (-\frac{3}{130} * \log_2(\frac{3}{130})) = 0.1583$

Entropy (Total , 1) $= (-\frac{1}{1} * \log_2(\frac{1}{1})) + (-\frac{0}{1} * \log_2(\frac{0}{1})) = 0$

Hasil perhitungan *entropy* lainnya dapat dilihat pada tabel 4.5

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.

3. Menghitung Nilai Information Gain

Setelah mendapatkan nilai *entropy* untuk suatu kumpulan sampel data, maka dapat diukur efektivitas suatu fitur dalam mengklasifikasikan data. Ukuran efektivitas ini disebut sebagai *information gain*. Perhitungan *information gain* menggunakan persamaan 2.2, berikut langkah – langkah perhitungan *information gain* :

$$IG\ (S,A) = \ Entropy(S)\ - \sum_{v \equiv Values(A)} \frac{|S_v|}{S} * Entropy(S_v)$$

Gain (Total, *duration*) = 0.7219 - $\left(\frac{189}{190} * 0.8487\right) + \left(\frac{1}{190} * 0\right) = 0.105$

Gain (Total, *src_bytes*) = 0.7219 - $\left(\frac{90}{99} * 0\right) + \left(\frac{9}{99} * 0\right) = 0.7219$

Gain (Total, *dst_bytes*) = 0.7219 - $\left(\frac{130}{131} * 0.1583\right) + \left(\frac{1}{131} * 0\right) = 0.5648$

4. Menghitung Symmetrical Uncertainty

Kemudian dilanjutkan dengan menghitung kompensasi bias *IG* terhadap fitur dengan nilai lebih tersendiri dan menormalkan nilai-nilai ke kisaran 0 hingga 1. Pengukuran *SU* dapat dihitung dengan persamaan 2.3 sebagai berikut :

$$SU\ (S,A) = \ 2 * \frac{IG(S,A)}{H(S) + H(A)}$$

SU (Total, *duration*) = 2 * $\left(\frac{0.105}{0.7129+0.3653}\right) = 0.1931$

SU (Total, *src_bytes*) = 2 * $\left(\frac{0.7219}{0.7129+0.6978}\right) = 1.0169$

SU (Total, *dstc_bytes*) = 2 * $\left(\frac{0.5648}{0.7129+0.5309}\right) = 0.9016$

Berikut merupakan tabel hasil dari perhitungan metode FCBF :

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.

Tabel 4.7 Hasil Perhitungan FCBF

Nama Fitur	Nilai Fitur	Total Kasus	Jumlah Serangan	Jumlah Normal	Entropy	IG	Split	SU
duration	0	189	137	52	0.8487	0.105	0.3653	0.1931
	1	1	1	0	0			
protocol type	0	19	19	0	0	0.0412	0.642	0.0604
	0.3125	229	177	52	0.7729			
	1	12	12	0	0			
src_bytes	0	90	90	0	0	0.7219	0.6978	1.0169
	1	1	4	5	0,7219			
dst_bytes	0	130	127	3	0.1584	0.6427	0.5309	1.0260
	1	1	1	0	0			

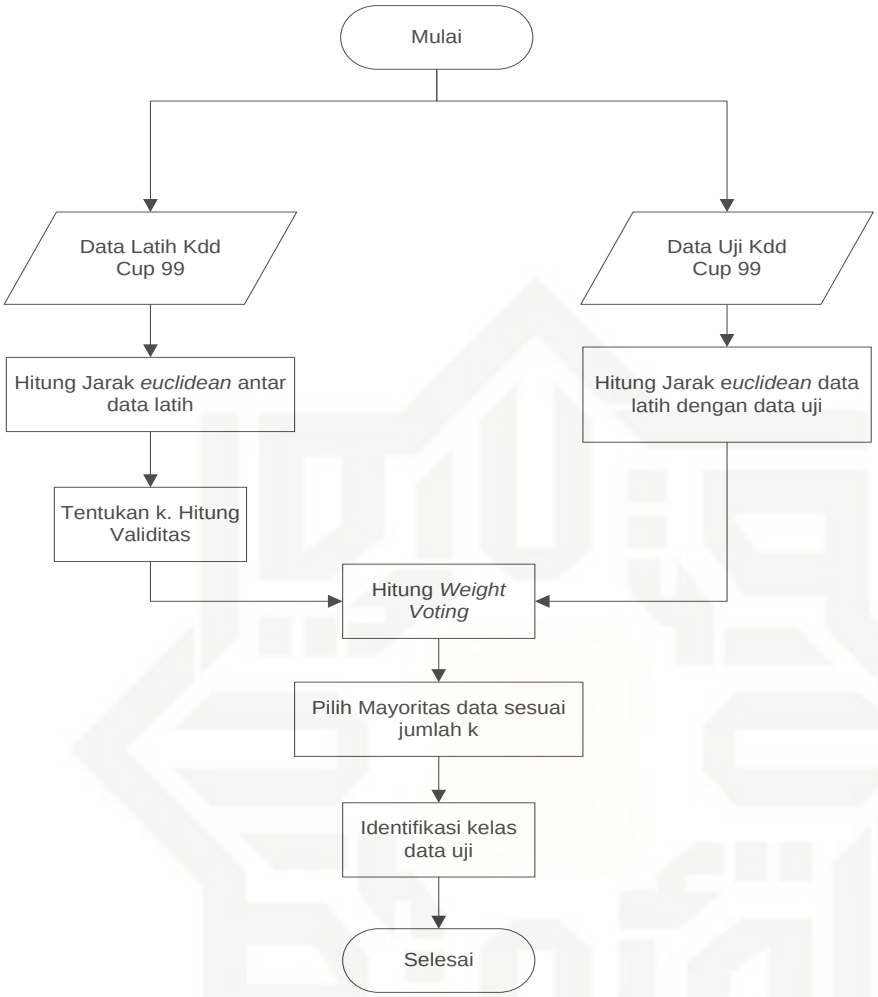
5. Menentukan pilihan fitur

Setelah mendapatkan nilai hasil dari SU selanjutnya dipilih fitur yang akan digunakan dalam klasifikasi MK-NN. Sehingga hasil yang diperoleh dan fitur yang akan digunakan adalah *duration, src_bytes, dst_bytes, wrong_fragment, count, same_srv_rate, srv_rerror_rate, dst_host_diff_srv_rrate, dst_host_srv_ count, dst_host_srv_serror_rate, dst_host_srv_rerror_rate, dst_host_srv_diff_ host_rate, dst_host_same_src_port_rate, hot, num_failed_lo gins, num_ compromised, root_sheel, su_attempted, num_root, num_file_crea tions, num_shells, num_access_files, num_outbound_cmds, is_host_login, is_guest_ login* dan *jenis*.

4.2.5 Klasifikasi dengan *Modified K-Nearest Neighbor* (MK-NN)

Berdasarkan data yang telah didapatkan pada proses sebelumnya, maka pada bagian ini dijelaskan bagaimana penggunaan metode MK-NN dalam klasifikasi data tersebut. Untuk lebih jelas mengenai cara kerja algortima MK-NN ini akan dijelaskan pada *flowchart* yang diperlihatkan pada gambar 4.2 dibawah ini :

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.
© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau



Gambar 4.2 Flowchart Klasifikasi *Modified K-Nearest Neighbor*

Berikut keterangan dari *flowchart* klasifikasi KDD CUP 99 dengan menggunakan metode *Modified K-Nearest Neighbor* :

1. Pembagian data latih dan data uji

Data latih merupakan semua data KDD CUP 99 yang digunakan dalam penelitian. Data latih yang digunakan adalah data latih yang telah melewati proses *selection, preprocessing, transformation* dan *fast correlation based filter*.

Data uji merupakan data yang akan ditentukan kelas klasifikasi dengan data latih yang telah ada di *database.* Pembagian data latih dan data uji dengan keadaan kelas data proporsional artinya ada kelas yang lebih mendominasi. Data uji yang digunakan pada penelitian ini dapat dilihat pada tabel 4.7 berikut :

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.

Tabel 4.7 Data Uji

no_id	duration	protocol _type	Land	wrong fragment	count	....	kelas
8	0	6	0	0	0	....	normal
75	0	6	0	0	0	....	dos

Tabel data uji diatas belum dilakukan proses normalisasi. Untuk tabel data uji yang sudah dilakukan normalisasi dapat dilihat pada table 4.8 dibawah ini:

Tabel 4.8 Data Uji yang Telah Dinormalisasi

no_id	duration	protocol _type	land	wrong fragment	count	....	Kelas
8	0	0.3125	0	0	0	....	normal
75	0	0.3125	0	0	0	....	dos

Tahap selanjutnya dilakukan proses perhitungan dengan menggunakan Mk-NN. Proses Mk-NN dibagi menjadi empat tahap perhitungan, yaitu jarak *Euclidean* antar data latih, *Validitas*, jarak *Euclidean* data latih dengan data uji dan *Weight Voting.*

2. Menghitung jarak *euclidean* data latih

Proses perhitungan jarak *euclidean* digunakan untuk mencari jarak *euclidean* antara data latih satu dengan data latih yang lain. Perhitungan jarak *euclidean* data latih menggunakan persamaan 2.2 untuk nilai data latih pada tabel 4.4, berikut langkah – langkah perhitungan jarak *euclidean* data latih :

a. Untuk jarak antara data latih 1 dan 2.

$$d\,(x,y)(datalatih\ x, data\ latih\ y) = \sqrt{\sum_{i=I}^{n}(x_i - y_i)^2}$$

$$d\,(1,2) = \sqrt{\begin{matrix}(0-0)^2+(0.3125-0.3125)^2+(0-0)^2+(0-0)^2+\\(0-0)^2+(0-0)^2+(0-0)^2+(1-1)^2\\+(0-0)^2+(0-0)^2+(0-0)^2+(0-0)^2+(0-0)^2+(0-0)^2\\+(0-0)^2+(0-0)^2+(0-0)^2+(0-0)^2+(0-0)^2+(0-0)^2\\+(1-1)^2+(0-0)^2+(0-0)^2+(1-1)^2+(0-0)^2+(0.11-0.05)^2\\+(0-0)^2+(0-0)^2+(0-0)^2+(0-0)^2+(0-0)^2\end{matrix}}$$

$$d\,(1,2) = \mathbf{0.06}$$

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.

b. Untuk jarak antara data latih 1 dan 3.

$$d\,(x,y)(datalatih\ x, data\ latih\ y) = \sqrt{\sum_{i=I}^{n}(x_i - y_i)^2}$$

$$d\,(1,3) = \sqrt{\begin{matrix}(0-0)^2 + (0.3125-0.3125)^2 + (0-0)^2 + (0-0)^2 + \\ (0-0)^2 + (0-0)^2 + (0-0)^2 + (1-1)^2 \\ +(0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 \\ (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 \\ +(0-0)^2 + (0-0)^2 + (1-1)^2 + (0-0)^2 + (0-0)^2 \\ +(1-1)^2 + (0-0)^2 + (0.11-0.03)^2 + (0-0)^2 + (0-0)^2 \\ +(0-0)^2 + (0-0)^2 + (0-0)^2\end{matrix}}$$

$$d\,(1,3) = \mathbf{0.08}$$

c. Untuk jarak antara data latih 2 dan 3.

$$d\,(x,y)(datalatih\ x, data\ latih\ y) = \sqrt{\sum_{i=I}^{n}(x_i - y_i)^2}$$

$$d\,(2,3) = \sqrt{\begin{matrix}(0-0)^2 + (0.3125-0.3125)^2 + (0-0)^2 + (0-0)^2 + \\ (0-0)^2 + (0-0)^2 + (0-0)^2 + (1-1)^2 \\ +(0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 \\ +(0-0)^2(0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 \\ +(0-0)^2 + (0-0)^2 + (1-1)^2 + (0-0)^2 + (0-0)^2 \\ +(1-1)^2 + (0-0)^2 + (0.05-0.03)^2 + (0-0)^2 \\ +(0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2\end{matrix}}$$

$$d\,(2,3) = \mathbf{0,02}$$

Perhitungan diatas dilakukan pada semua data latih untuk memperoleh jarak *euclidean* secara keseluruhan. Perhitungan dilakukan dengan cara dan persamaan yang sama. Nilai dari perhitungan jarak *euclidean* data latih dapat dilihat pada Tabel 4.9. Diketahui *P* adalah sampel data.

Tabel 4.9 Nilai Jarak Antar Data Latih

d	P1	P2	P3	P4	P5	P6	....	P260
P1	0						....	
P2	0,06	0					....	
P3	0,08	0,02	0				....	
P4	0,08	0,02	0	0			....	
P5	0,09	0,03	0,01	0,01	0		....	
P6	0,09	0,03	0,01	0,01	0	0	....	
P7	1,3392	1,3798	1,3937	1,3937	1,4007	1,4007	....	

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.

P9	0,0412	0,0806	0,0984	0,0984	0,1077	0,1077	….	
….	….	….	….	….	….	….	….	
P260	0,3923	0,4601	0,4936	0,5130	0,5464	0,5749	….	0

3. Menghitung nilai validitas data latih

Menghitung validitas pada setiap data latih menggunakan persamaan 2.3 pemberian nilai 1 atau 0 untuk setiap ketetanggan menggunakan persamaan 2.4. Sebelum melakukan perhitungan, tentukan nilai k atau jarak ketetanggaannya terlebih dahulu. Pada contoh perhitungan ini ditentukan nilai k = 3. mengacu pada tabel 4.7, berikut adalah langkah-langkah untuk mencari nilai validitasnya:

$$Validitas(\mathrm{x}) = \frac{1}{k}\sum_{i=1}^{k} S\,(label(x), (label(N_i(x))$$

$$Validitas\ (\mathrm{data}_1) = \frac{1}{3} * (1 + 1 + 1)$$

$$Validitas\ (\mathrm{data}_1) = 1$$

Penjelasan:

- $label(x)$= data x (dalam contoh di atas x = data ke-1)

- $label(N_i(x))$ = data dengan jarak euclidean terdekat dengan data x

- kemudian bandingkan kelas pada $label(x)$ dan kelas pada $label(N_i(x))$. Jika kedua kelas bernilai sama, maka beri nilai 1. Namun, jika kedua kelas bernilai beda, maka beri nilai 0.

- lakukan perbandingan kelas $label(x)$ dan kelas $label(N_i(x))$ sebanyak nilai k yang digunakan (contoh di atas k yang digunakan = 3).

Perhitungan diatas dilakukan pada semua data latih untuk memperoleh nilai validitas. Perhitungan dilakukan dengan cara dan persamaan yang sama. Nilai dari perhitungan validitas data latih dapat dilihat pada Tabel 4.10.

Tabel 4.10 Validitas Data Latih

Data ke-	Validitas
1	0.6667
2	1
3	1
4	1
5	1

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.

Data ke-	Validitas
6	1
7	1
8	1
9	1
…	…
234	0.3333

4. Menghitung jarak *euclidean* data uji dengan data latih

Setelah jarak *euclidean* antar data latih dihitung, selanjutnya menghitung jarak *euclidean* antar data uji dengan data latih. Untuk menghitung jarak *euclidean* data uji dengan data latih menggunakan persamaan 2.2 dan mengacu pada tabel 4.5 dan 4.7, berikut langkah perhitungan jarak *euclidean* data uji dengan data latih :

$$d(dataujix,\ datalatihx) = \sqrt{\sum_{i=I}^{n}(x_{2i} - x_{1i})^2}$$

$$\text{d(datauji1, datalatih1)} = \sqrt{\begin{matrix}(0-0)^2 + (0.3125-0.3125)^2 + (0-0)^2 \\ +(0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 \\ +(1-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 \\ +(0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 \\ +(0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 \\ +(0-0)^2 + (1-1)^2 + (0-0)^2 + (0-0)^2 \\ +(1-1)^2 + (0-0)^2 + (0.09-0.11)^2 + (0.04-0)^2 \\ +(0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2\end{matrix}}$$

$$= \mathbf{0.0447}$$

$$\text{d(datauji2, datalatih1)} = \sqrt{\begin{matrix}(0-0)^2 + (0.3125-0.3125)^2 + (0-0)^2 \\ +(0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 \\ +(0-1)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 \\ +(0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 \\ +(0-0)^2 + (1-0)^2 + (1-0)^2 + (0-0)^2 \\ +(0-0)^2 + (0.1313-1)^2 + (0.06-0)^2 + (0-0)^2 \\ +(0.0563-1)^2 + (0.14-0)^2 + (0.05-0.11)^2 \\ +(0-0)^2 + (1-0)^2 + (1-0)^2 + (0-0)^2 + (0-0)^2\end{matrix}}$$

$$= \mathbf{2.583}$$

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.

Perhitungan diatas dilakukan pada data latih dengan data uji untuk memperoleh jarak *euclidean*. Perhitungan dilakukan dengan cara dan persamaan yang sama. Nilai dari perhitungan jarak *euclidean* data uji dengan data latih dapat dilihat pada Tabel 4.11.

Tabel 4.11 Nilai Jarak *Euclidean* Data Uji dengan Data Latih

No	d_e data uji 1	d_e data uji 2
1	0,0447	2,583
2	0,0565	2,5823
3	0,0721	2,5824
4	0,0721	2,5824
5	0,0806	2,5825
6	0,0806	2,5825
7	1,3520	2,9278
8	0,03	2,5835
9	0,0639	2,5843
…	…	…
234	1,2815	2,5186

5. Menghitung weight voting

Weight voting digunakan untuk menentukan kelas dari data uji. Nilai *weight voting* didapatkan dari perhitungan validitas dan jarak *euclidean* antara data uji dengan data latih. Setelah hasil perhitungan diperoleh, nilai *weight voting* yang bernilai paling besarlah yang akan digunakan sebagai penentu kelas. Perhitungan *weight voting* menggunakan persamaan 2.6, mengacu pada tabel 4.10 dan tabel 4.11, berikut langkah – langkah perhitungan *weight voting* antara data uji dengan data latih :

$$\mathrm{W}(i) = Validitas(i) * \frac{1}{d_e + 0{,}5}$$

$$\mathrm{W}(\mathrm{datauji1}) = 0.6667 * \frac{1}{0.0447 + 0.5} = 1.2238$$

$$\mathrm{W}(\mathrm{datauji2}) = 0.6667 * \frac{1}{2.5830 + 0.5} = 0.2162$$

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.

Perhitungan diatas dilakukan pada data uji dengan data latih untuk memperoleh nilai *weight voting*. Perhitungan dilakukan dengan cara dan persamaan yang sama. Nilai dari perhitungan *weight voting* data uji dengan data latih dapat dilihat pada Tabel 4.12.

Tabel 4.12 Nilai *Weight Voting*

No	WV data uji 1	WV data uji 2
1	1,2238	0,2162
2	1,7967	0,3244
3	1,7479	0,3244
4	1,7479	0,3244
5	1,7222	0,3244
6	1,7222	0,3244
7	0,5399	0,2917
8	1,8867	0,3242
9	1,7731	0,3242
…	…	…
234	0,1871	0,1104

6. Menentukan kelas data uji berdasarkan nilai k

Nilai *weight voting* yang telah diperoleh digunakan untuk menentukan kelas dari data uji. Nilai *weight voting* tersebut diurutkan dari yang terbesar hingga terkecil kemudian ambil sebanyak *k* yang telah ditentukan. Nilai *weight voting* yang telah diurutkan dari total 260 data yang telah dilakukan perhitungan dapat dilihat pada Tabel 4.13.

Tabel 4.13 Urutan Nilai *Weight Voting* dari yang Terbesar Hingga Terkecil

No	WV data uji 1	WV data uji 2
1	1.9607	1.6540
2	1.9449	1.5304
3	1.9449	1.5085

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.

© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau
UIN SUSKA RIAU

Setelah hasil k tertinggi dari weight voting didapat, maka cari kelas dari setiap data weight voting tertinggi. Kelas asli dari weight voting dan mayoritasnya dapat dilihat pada tabel 4.14 dibawah ini :

Tabel 4.14 Kelas Asli Hasil Weight Voting

No	WV data uji 1	WV data uji 2
1	Normal	Dos
2	Normal	Dos
3	Normal	Dos
Mayoritas	Normal	Dos

Setelah didapat mayoritas kelas, maka hasil klasifikasi dibandingkan dengan kelas asli data uji. maka didapatlah akurasi kecocokan antara kelas yang diprediksi dan kelas pada data sebenarnya. Hasil klasifikasi dapat dilihat pada table 4.15 dibawah ini :

Tabel 4.15 Hasil Klasifikasi

No	Kelas Asli	Kelas Hasil Klasifikasi	Terkenali
1	Normal	Normal	Benar
2	Dos	Dos	Benar

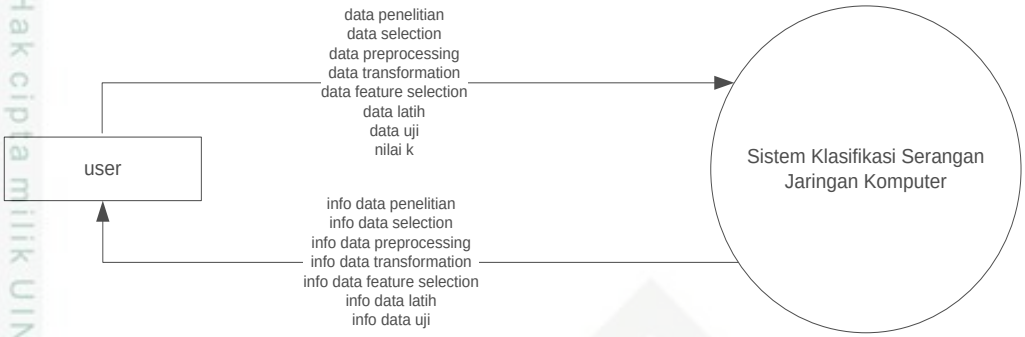
4.3 Analisa Sistem

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

4.3.1 *Context Diagram*

Context Diagram menggambarkan aliran fungsional dalam sebuah proses pada sistem. *Context Diagram* akan dijelaskan pada Gambar 4.3.

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.



Gambar 4.3 Context Diagram

Pada Gambar 4.3 *Context Diagram* memiliki entitas yakni pengguna *user*. Aliran data terjadi antara entitas tersebut dengan sistem. Data yang dialirkan dari *User* ke sistem adalah data penelitian, data *selection*, data *preprocessing*, data *transformation*, data *feature selection*, data latih, data uji dan nilai *K*. Umpan balik atau data yang telah diproses sistem dan dialirkan kepada *user* adalah info data *selection*, info data *preprocessing*, info data *transformation*, info data *feature selection*, info data latih, info data uji.

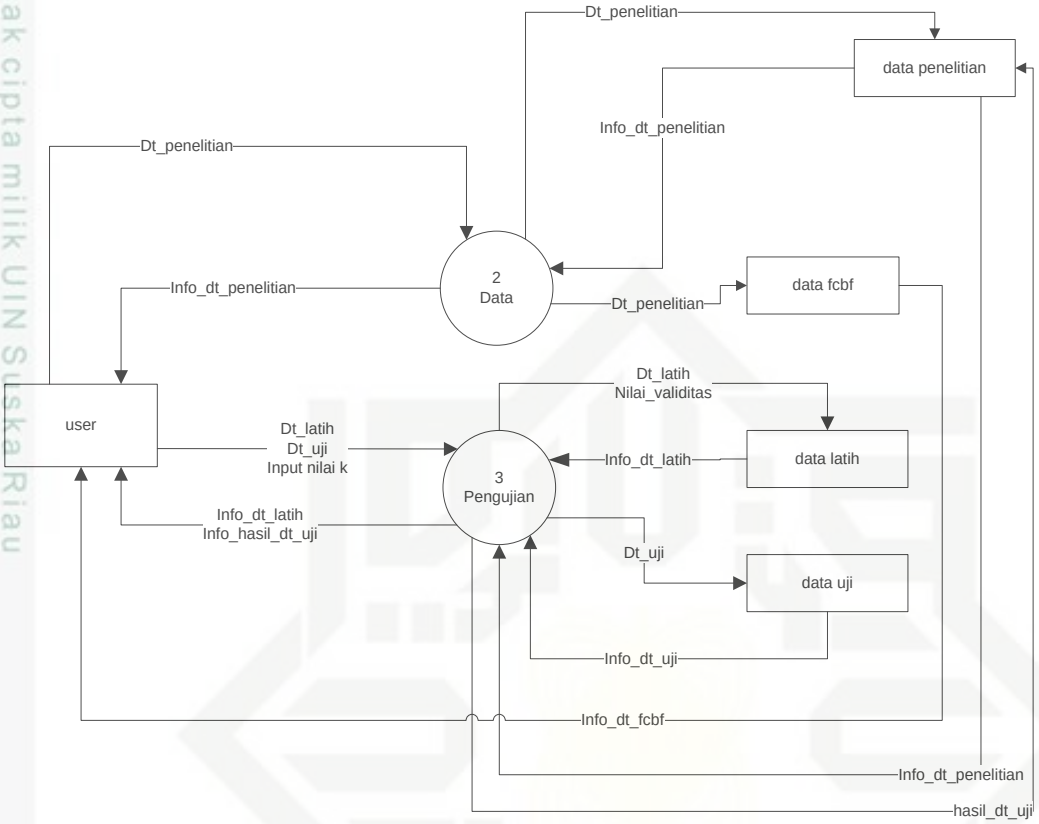
4.3.2 *Data Flow Diagram (DFD)*

Data Flow Diagram (DFD) merupakan diagram yang menggunakan notasi simbol untuk menggambarkan arus data sistem. Proses kerja sistem dapat dilihat pada Data Flow Diagram pada Gambar 4.4.

1. *Data Flow Diagram* (DFD) Level 1

Gambar 4.4 dibawah ini adalah gambaran DFD level 1 dari sistem klasifikasi kelas jenis serangan pada dataset KDD CUP 99.

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.
© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau



Gambar 4.4 *Data Flow Diagram* level 1

Pada DFD level 1 ada 3 proses yaitu proses *Login*, Data dan proses Pengujian. Untuk lebih jelasnya dapat dilihat pada Tabel 4.16

Tabel 4.16 Proses DFD Level 1

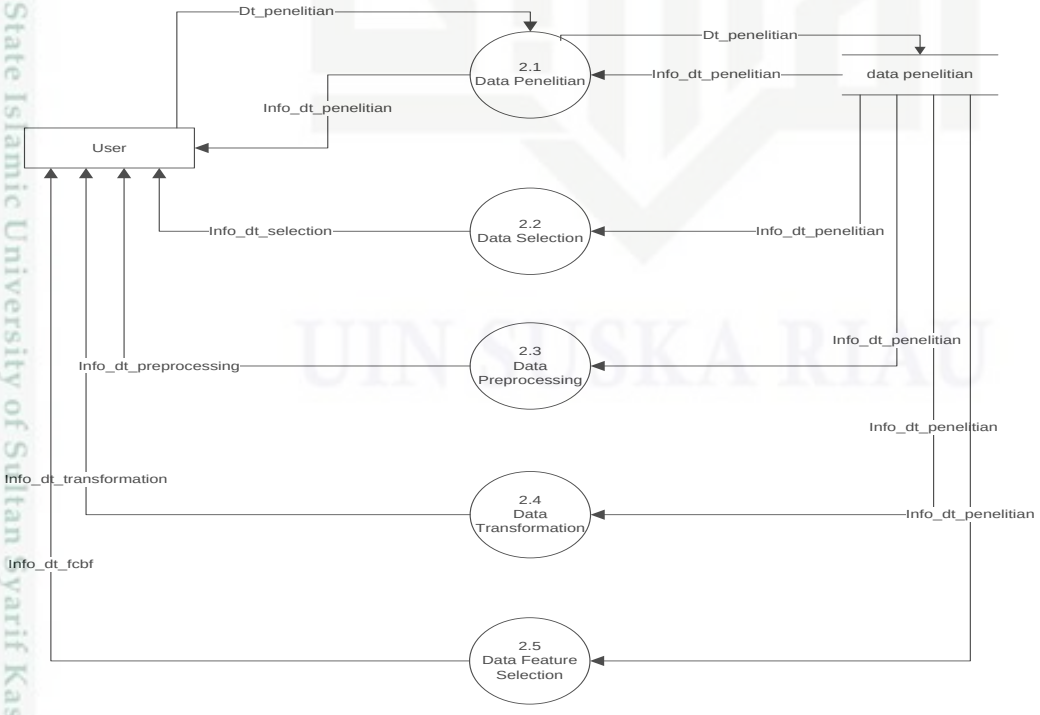
No	Proses	Deskripsi
1	Data	- Terjadi *input*-an data penelitian oleh *user*. - Data penelitian yang di *input* kan diproses dan disimpan kedalam *database*. - *User* mendapat umpan balik berupa data penelitian.
2	Pengujian	- Terjadi *input*-an data latih, data uji dan nilai k oleh *user*. - Data penelitian dan data fcbf yang ada dalam *database* kemudian diproses menjadi data latih dan data uji. - Setelah diproses kemudian data latih dan nilai validitas

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.
© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau

No	Proses	Deskripsi
		disimpan kedalam tabel data latih dan data uji disimpan kedalam tabel data uji. - *User* mendapat umpan balik berupa data latih dan data uji. - Selanjutnya terjadi *input*-an data uji oleh *user*. - *Input*-an yang berasal dari *user* merupakan data uji dan seluruh data penelitian menjadi data latih kemudian diproses untuk mendapatkan hasil klasifikasi. - *User* menerima umpan balik berupa hasil klasifikasi data uji. - Kemudian hasil klasifikasi data uji disimpan kedalam tabel data penelitian sebagai data pembelajaran baru.

2. *Data Flow Diagram* (DFD) Level 2 proses 2 (Data)

Gambar 4.5 dibawah ini adalah gambaran DFD level 2 proses 2 dari sistem klasifikasi serangan jaringan komputer.



Gambar 4.5 *Data Flow Diagram* level 2 proses 2

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.
© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau

Pada DFD level 2 proses 2 ini terdiri dari proses Data Penelitian, Preprocessing, Transformasi dan Seleksi Fitur. Untuk lebih jelasnya dapat dilihat pada Tabel 4.17

Tabel 4.17 Proses DFD Level 2 Proses 2

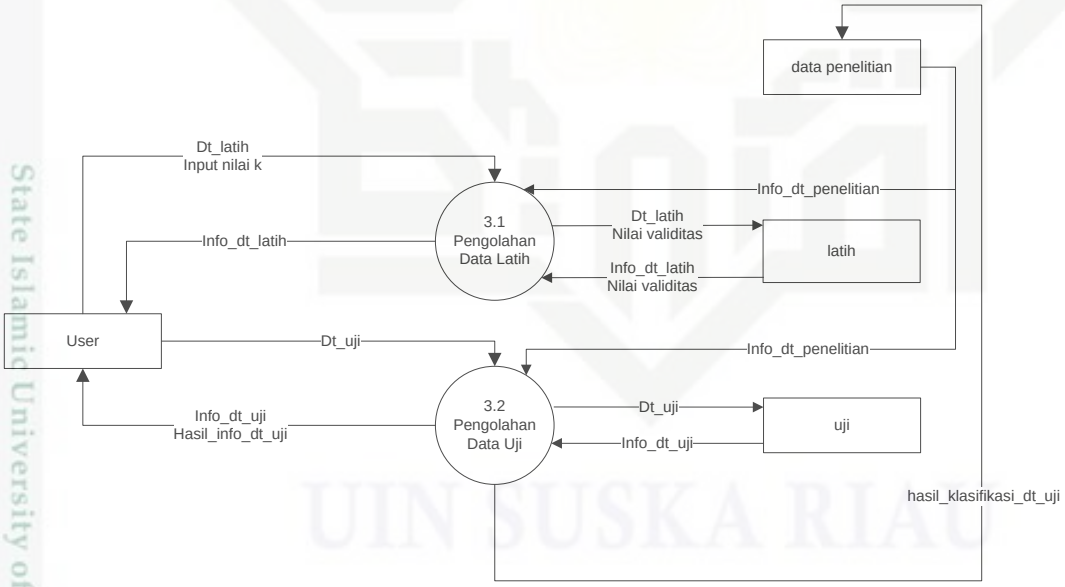
No	Proses	Deskripsi
1	Data Penelitian	- Terjadi *input*-an data penelitian oleh *user*. - Data penelitian yang di-*input*-kan diproses dan disimpan kedalam *database*. - *User* mendapat umpan balik berupa data penelitian.
2	*Data Selection*	- Terjadi *input*-an data penelitian oleh *user*. - Data penelitian yang di-*input*-kan dilakukan proses seleksi dengan menghapus fitur-fitur yang tidak diperlukan - *User* mendapat umpan balik berupa data penelitian yang telah diseleksi.
3	*Data Proprocessing*	- Terjadi *input*-an data penelitian oleh *User*. - Data penelitian yang di-*input*-kan dilakukan proses *preprocessing* dengan menggunakan *cleaning* atau pembersihan dengan cara menghapus data yang redudan. - *User* mendapat umpan balik berupa data penelitian yang telah dilakukan *preprocessing*.
4	*Data Transformation*	- Terjadi *input*-an data penelitian oleh *user*. - Data penelitian yang di-*input*-kan dilakukan proses transformasi dengan mengubah bentuk data penelitian menjadi normalisasi dengan mengubah nila *range* data berada pada rentang [0-1]. - *User* mendapat umpan balik berupa data

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.
© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau
UIN SUSKA RIAU

No	Proses	Deskripsi
		penelitian yang telah ditransformasi.
5	*Data Feature Selection*	- Terjadi *input*-an data penelitian oleh *user*. - Data penelitian yang di-*input*-kan dilakukan proses *feature selecton* menggunakan metode *fast correlation based filter* dengan memilih dan me-revisi fitur-fitur pada data penelitian yang akan digunakan selanjutnya pada proses klasifikasi - *User* mendapat umpan balik berupa data penelitian yang telah ditransformasi.

3. *Data Flow Diagram* (DFD) Level 2 proses 3 (Pengujian)

Gambar 4.6 dibawah ini adalah gambaran DFD level 2 proses 3 dari sistem klasifikasi serangan jaringan komputer.



Gambar 4.5 *Data Flow Diagram* level 2 proses 3

Pada DFD level 2 proses 3 ini terdiri dari proses Pengolahan Data Latih dan Pengolahan Data Uji. Untuk lebih jelasnya dapat dilihat pada Tabel 4.18

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.
© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau
UIN SUSKA RIAU

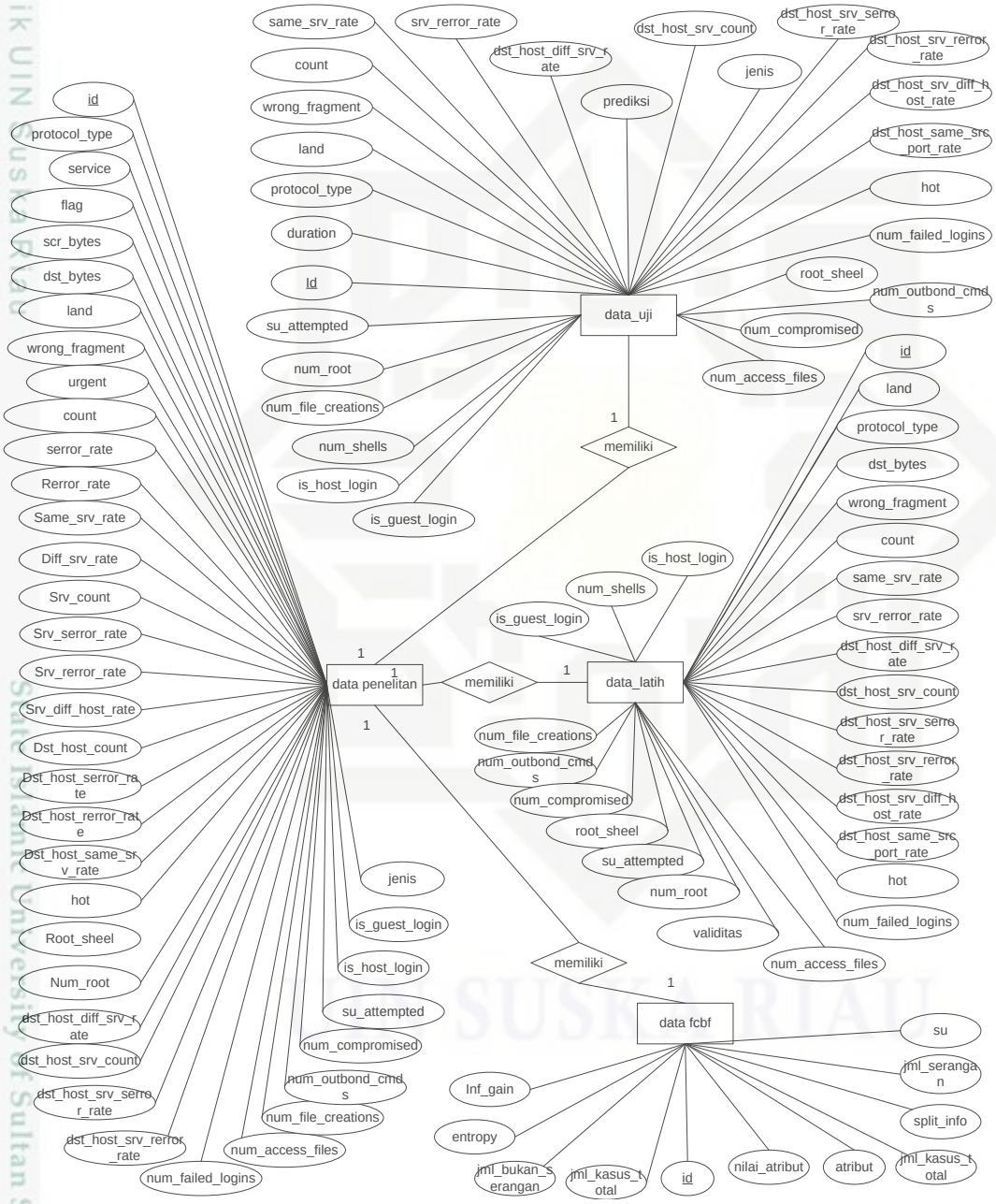
Tabel 4.18 Proses DFD Level 2 Proses 3

No	Proses	Deskripsi
1	Pengolahan Data Latih	- Terjadi *input*-an data latih dan nilai k oleh *user*. - Data penelitian yang ada dalam *database* kemudian diproses untuk mengolah data latih. - Data latih dan nilai validitas yang diperoleh setelah dilkukan proses perhitungan sesuai dengan nilai k yang di-*input*-kan kemudian disimpan kedalam *database*. - *User* mendapat umpan balik berupa data latih yang sudah diproses.
2	Pengolahan Data Uji	- Terjadi *input*-an data uji oleh *user*. - Data penelitian yang ada dalam *database* kemudian diproses untuk mengolah data uji. - Data uji kemudian diproses dan disimpan kedalam *database*. - *User* mendapat umpan balik berupa data hasil klasifikasi data uji. - Selanjutnya terjadi *input*-an data uji oleh *user*. - *input*-an yang berasal dari *user* kemudian diproses untuk mendapatkan hasil klasifikasi. - *user* menerima umpan balik berupa hasil klasifikasi data uji. - Kemudian hasil klasifikasi data uji disimpan kedalam tabel data penelitian sebagai data pembelajaran baru.

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.

4.3.3 *Entity Relationship Diagram (ERD)*

ERD memperlihatkan entitas-entitas yang terlibat dalam sebuah sistem serta relasi antar entitas tersebut. ERD pada penelitian ini dapat dilihat pada Gambar 4.6.



Gambar 4.6 ***Entity Relationship Diagram (ERD)***

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.
© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau
UIN SUSKA RIAU

Keterangan dari *Entity Relational Diagram (ERD)* dapat dilihat pada tabel 4.19 di bawah ini.

Tabel 4.19 Keterangan Entity relationship Diagram (ERD)

No	Nama	Deskripsi	Atribut	*Primary Key*	*Foreign Key*
1	Data Penelitian	Tabel untuk menyimpan data penelitian	- id - duration - protocol_tipe - service - flag - src_bytes - dst_bytes - land - wrong_fragment - urgent - count - serror_rate - rerror_rate - same_srv_rate - diff_srv_rate - srv_count - srv_serror_rate - srv_rerror_rate - dst_host_diff_srv_rate - dst_host_srv_count - dst_host_srv_serror_rate - dst_host_srv_rerror_rate - dst_host_same_srv_rate - dst_host_diff_srv_rate - dst_host_srv_count - dst_host_srv_serror_rate	id	-

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.
© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau
UIN SUSKA RIAU

No	Nama	Deskripsi	Atribut	*Primary Key*	*Foreign Key*
			- dst_host_srv_rerror_rate - dst_host_srv_diff_host_rate - dst_host_same_src_port_rate - hot - num_failed_logins - logged_in - num_compromised - root_sheel - su_attempted - num_root - num_file_creations - num_shells - num_access_files - num_outbound_cmds - is_host_login - is_guest_login - jenis		
2	FCBF	Tabel untuk menyimpan seleksi fitur	- id - atribut - nilai_atribut - jml_kasus_total - jml_serangan - jml_bukan_serangan - entropy - inf_gain - split_info - su	id	
3	Latih	Tabel untuk	- id - duration	id	

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.
© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau
UIN SUSKA RIAU

No	Nama	Deskripsi	Atribut	*Primary Key*	*Foreign Key*
		menyimpan data latih	- protocol_type - land - wrong_fragment - urgent - count - serror_rate - rerror_rate - same_srv_rate - srv_serror_rate - srv_rerror_rate - srv_diff_host_rate - dst_host_count - dst_host_srv_serror_rate - dst_host_srv_rerror_rate - dst_host_srv_diff_host_rate - dst_host_same_src_port_rate - hot - num_failed_logins - logged_in - su_attempted - num_root - num_file_creations - num_shells - num_access_files - num_outbound_cmds - is_host_login - is_guest_login - diff_srv_rate - srv_count		

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.
© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau
UIN SUSKA RIAU

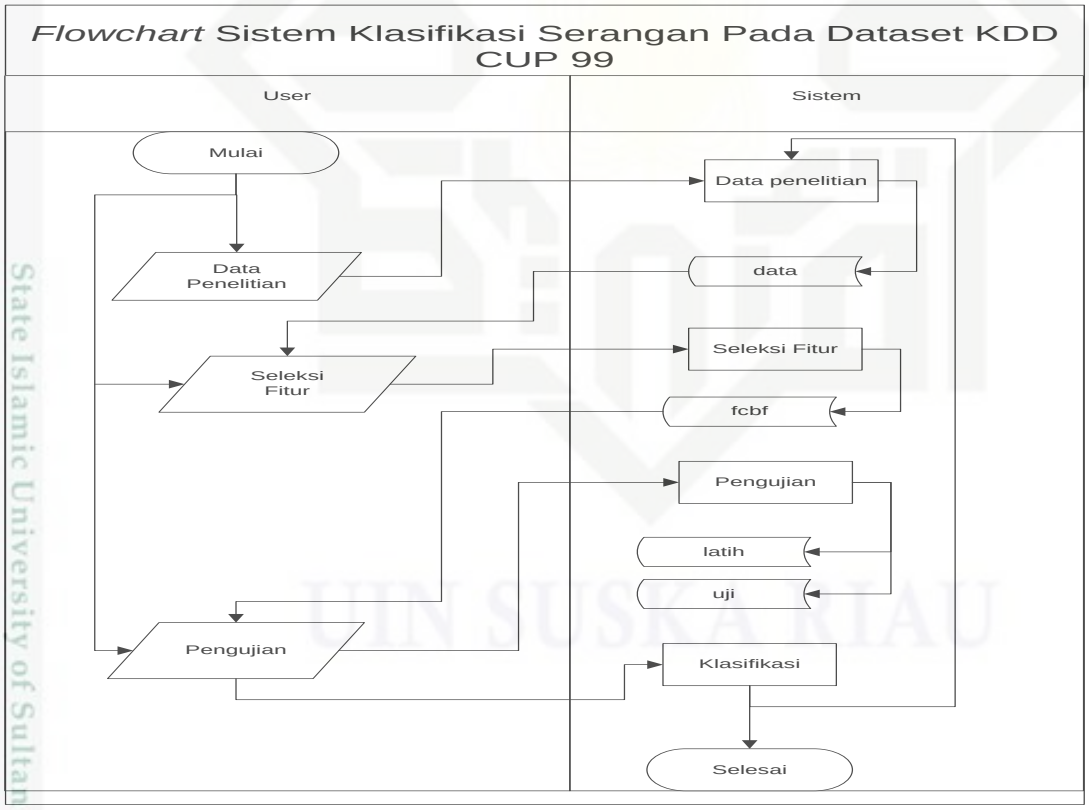
No	Nama	Deskripsi	Atribut	*Primary Key*	*Foreign Key*
			- dst_host_same_srv_rate - jenis - validitas		
4	Uji	Tabel untuk menyimpan data uji	- id - duration - protocol_type - land - wrong_fragment - urgent - count - serror_rate - rerror_rate - same_srv_rate - srv_serror_rate - srv_rerror_rate - srv_diff_host_rate - dst_host_count - dst_host_srv_serror_rate - dst_host_srv_rerror_rate - dst_host_srv_diff_host_rate - dst_host_same_src_port_rate - hot - num_failed_logins - logged_in - su_attempted - num_root - num_file_creations - num_shells - num_access_files	id	-

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.
© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau
UIN SUSKA RIAU

No	Nama	Deskripsi	Atribut	*Primary Key*	*Foreign Key*
			- num_outbound_cmds - is_host_login - is_guest_login - diff_srv_rate - srv_count - dst_host_same_srv_rate - jenis - prediksi		

4.3.4 *Flowchart*

Flowchart sistem klasifikasi ini dapat dilihat pada Gambar 4.8.



Gambar 4.8 Flowchart Sistem Klasifikasi Serangan Pada Dataset KDD CUP 99

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.

4.4 Perancangan Basis Data (*Database*)

Berikut adalah perancangan tabel *database* pada penelitian ini.

4.4.1 Tabel Data Penelitian

Nama tabel : data_penelitian

Deskripsi isi : Berisi data penelitian yang digunakan sebagai pembelajaran untuk pengujian sistem.

Primary key : ***id***

Penjelasan struktur tabel akses dapat dilihat selengkapnya pada Tabel 4.20.

Tabel 4.20 Struktur Tabel Data Penelitian

No	Nama Kolom	Tipe Dan Panjang Data	Null	Keterangan
1	id	Int (5)	Not Null	Primery key
2	duration	Varchar (100)	Not Null	
2	protocol_tipe	Varchar (100)	Not Null	
3	service	Varchar (100)	Not Null	
4	flag	Varchar (100)	Not Null	
5	src_bytes	Varchar (100)	Not Null	
6	dst_bytes	Varchar (100)	Not Null	
7	land	Varchar (100)	Not Null	
8	wrong_fragment	Varchar (100)	Not Null	
9	urgent	Varchar (100)	Not Null	
10	count	Varchar (100)	Not Null	
11	serror_rate	Varchar (100)	Not Null	
12	rerror_rate	Varchar (100)	Not Null	
13	same_srv_rate	Varchar (100)	Not Null	
14	diff_srv_rate	Varchar (100)	Not Null	
15	srv_count	Varchar (100)	Not Null	
16	srv_serror_rate	Varchar (100)	Not Null	
17	srv_rerror_rate	Varchar (100)	Not Null	
18	dst_host_diff_srv_rate	Varchar (100)	Not Null	

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.

No	Nama Kolom	Tipe Dan Panjang Data	Null	Keterangan
19	dst_host_srv_count	Varchar (100)	Not Null	
20	dst_host_srv_serror_rate	Varchar (100)	Not Null	
21	dst_host_srv_rerror_rate	Varchar (100)	Not Null	
22	dst_host_same_srv_rate	Varchar (100)	Not Null	
23	dst_host_diff_srv_rate	Varchar (100)	Not Null	
24	dst_host_srv_count	Varchar (100)	Not Null	
25	dst_host_srv_serror_rate	Varchar (100)	Not Null	
26	dst_host_srv_rerror_rate	Varchar (100)	Not Null	
27	dst_host_srv_diff_host_rate	Varchar (100)	Not Null	
28	dst_host_same_src_port_rate	Varchar (100)	Not Null	
29	hot	Varchar (100)	Not Null	
30	num_failed_logins	Varchar (100)	Not Null	
31	logged_in	Varchar (100)	Not Null	
32	num_compromised	Varchar (100)	Not Null	
33	root_sheel	Varchar (100)	Not Null	
34	su_attempted	Varchar (100)	Not Null	
35	num_root	Varchar (100)	Not Null	
36	num_file_creations	Varchar (100)	Not Null	
37	num_shells	Varchar (100)	Not Null	
38	num_access_files	Varchar (100)	Not Null	
39	num_outbound_cmds	Varchar (100)	Not Null	
40	is_host_login	Varchar (100)	Not Null	
41	is_guest_login	Varchar (100)	Not Null	
42	jenis	Varchar (100)	Not Null	

4.4.2 Tabel Data Latih

Nama tabel : latih

Deskripsi isi : Berisi data latih untuk pengujian sistem.

Primary key : ***id***

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.

Penjelasan struktur tabel data_latih dapat dilihat selengkapnya pada Tabel 4.21.

Tabel 4.21 Struktur Tabel Data Latih

No	Nama Kolom	Tipe Dan Panjang Data	Null	Keterangan
1	id	Int (5)	Not Null	Primery key
2	duration	Varchar (100)	Not Null	
3	protocol_type	Varchar (100)	Not Null	
4	land	Varchar (100)	Not Null	
5	wrong_fragment	Varchar (100)	Not Null	
6	count	Varchar (100)	Not Null	
7	same_srv_rate	Varchar (100)	Not Null	
8	srv_rerror_rate	Varchar (100)	Not Null	
9	dst_host_diff_srv_rate	Varchar (100)	Not Null	
10	dst_host_srv_count	Varchar (100)	Not Null	
11	dst_host_srv_serror_rate	Varchar (100)	Not Null	
12	dst_host_srv_rerror_rate	Varchar (100)	Not Null	
13	dst_host_srv_diff_host_rate	Varchar (100)	Not Null	
14	dst_host_same_src_port_rate	Varchar (100)	Not Null	
15	hot	Varchar (100)	Not Null	
16	num_failed_logins	Varchar (100)	Not Null	
17	num_compromised	Varchar (100)	Not Null	
18	root_sheel	Varchar (100)	Not Null	
19	su_attempted	Varchar (100)	Not Null	
20	num_root	Varchar (100)	Not Null	
21	num_file_creations	Varchar (100)	Not Null	
22	num_shells	Varchar (100)	Not Null	
23	num_access_files	Varchar (100)	Not Null	
24	num_outbound_cmds	Varchar (100)	Not Null	
25	is_host_login	Varchar (100)	Not Null	
26	is_guest_login	Varchar (100)	Not Null	

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.

No	Nama Kolom	Tipe Dan Panjang Data	Null	Keterangan
27	jenis	Varchar (100)	Not Null	
28	Validitas	*Float*	Not Null	

4.4.3 Tabel Data Uji

Nama tabel : uji

Deskripsi isi : Berisi data uji untuk pengujian sistem.

Primary key : ***id***

Penjelasan struktur tabel data_uji dapat dilihat selengkapnya pada Tabel 4.22.

Tabel 4.22 Struktur Tabel Data Uji

No	Nama Kolom	Tipe Dan Panjang Data	Null	Keterangan
1	id	Int (5)	Not Null	Primery key
2	duration	Varchar (100)	Not Null	
3	protocol_type	Varchar (100)	Not Null	
4	land	Varchar (100)	Not Null	
5	wrong_fragment	Varchar (100)	Not Null	
6	count	Varchar (100)	Not Null	
7	same_srv_rate	Varchar (100)	Not Null	
8	srv_rerror_rate	Varchar (100)	Not Null	
9	dst_host_diff_srv_rate	Varchar (100)	Not Null	
10	dst_host_srv_count	Varchar (100)	Not Null	
11	dst_host_srv_serror_rate	Varchar (100)	Not Null	
12	dst_host_srv_rerror_rate	Varchar (100)	Not Null	
13	dst_host_srv_diff_host_rate	Varchar (100)	Not Null	
14	dst_host_same_src_port_rate	Varchar (100)	Not Null	
15	hot	Varchar (100)	Not Null	
16	num_failed_logins	Varchar (100)	Not Null	
17	num_compromised	Varchar (100)	Not Null	
18	root_sheel	Varchar (100)	Not Null	

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.

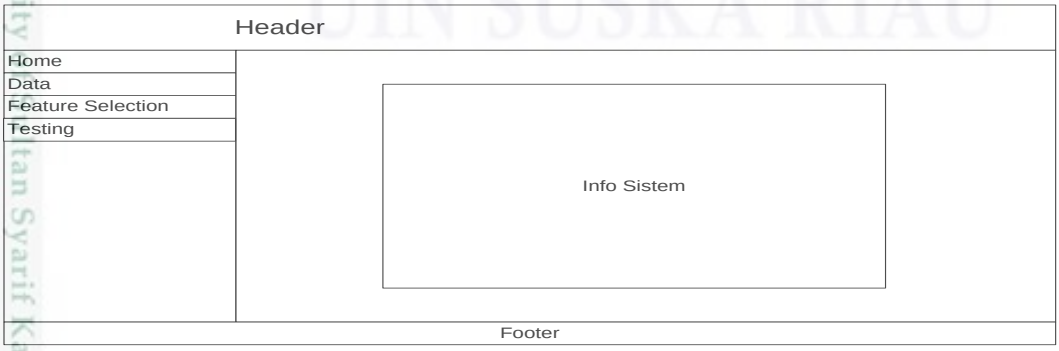
No	Nama Kolom	Tipe Dan Panjang Data	Null	Keterangan
19	su_attempted	Varchar (100)	Not Null	
20	num_root	Varchar (100)	Not Null	
21	num_file_creations	Varchar (100)	Not Null	
22	num_shells	Varchar (100)	Not Null	
23	num_access_files	Varchar (100)	Not Null	
24	num_outbound_cmds	Varchar (100)	Not Null	
25	is_host_login	Varchar (100)	Not Null	
26	is_guest_login	Varchar (100)	Not Null	
27	jenis	Varchar (100)	Not Null	
28	klasifikasi	Varchar (100)	Not Null	

4.5 Perancangan Antarmuka (*Interface*)

Antar muka (*Interface*) sistem merupakan sebuah sarana pengembangan sistem yang digunakan untuk membuat komunikasi dan penyampaian informasi lebih mudah dimengerti. *Interface* meliputi tampilan yang baik, mudah dipahami serta tombol - tombol yang *familiar*.

4.5.1 Perancangan Halaman *Home*

Halaman *home* adalah halaman paling awal ketika pengguna mengakses sistem. Secara umum, isi dari halaman beranda ini adalah penjelasan tentang menu-menu yang dapat diakses pengguna. Perancangan halaman beranda dapat dilihat pada Gambar 4.7.

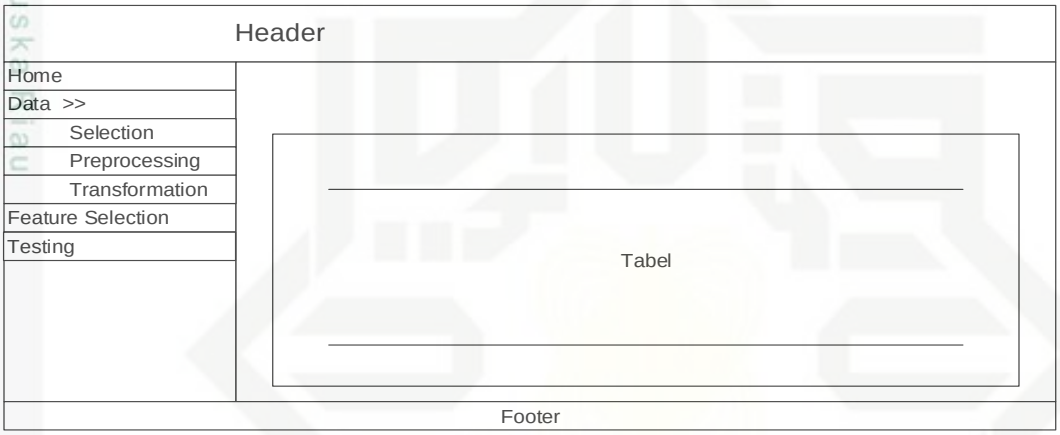


Gambar 4.7 Perancangan Halaman *Home*

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.

4.5.2 Perancangan Halaman Data

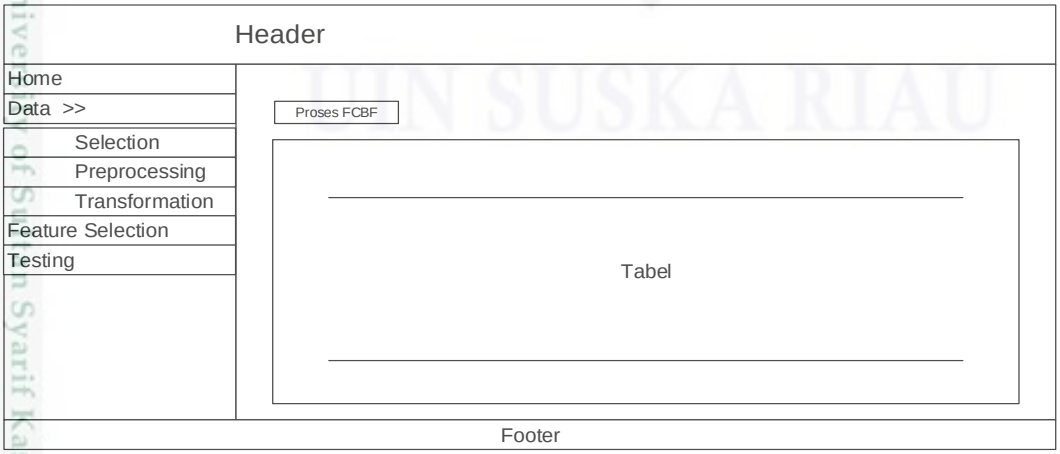
Halaman data ini merupakan halaman yang mempunyai *submenu* dan diakses oleh *user* untuk mengelola data *selection*, *preprocessing* dan *transformation* yang digunakan dalam sistem. Pada perancangan halaman data *selection*, *preprocessing* dan *transformation* mempunyai tampilan yg sama karena hanya menampilkan berupa data. Perancangan halaman dapat dilihat pada Gambar 4.8.



Gambar 4.8 Perancangan Halaman *Selection*, *Preprocessing* dan *Transformation*

4.5.3 Perancangan Halaman *Feature Selection*

Halaman seleksi fitur ini merupakan halaman yang dapat diakses oleh *user* yang akan melakukan proses seleksi fitur terhadap fitur-fitur yang terdapat pada KDD CUP 99 dengan menggunakan metode *Fast Correlation Based Filter*. Perancangan halaman ini dapat dilihat pada Gambar 4.9.

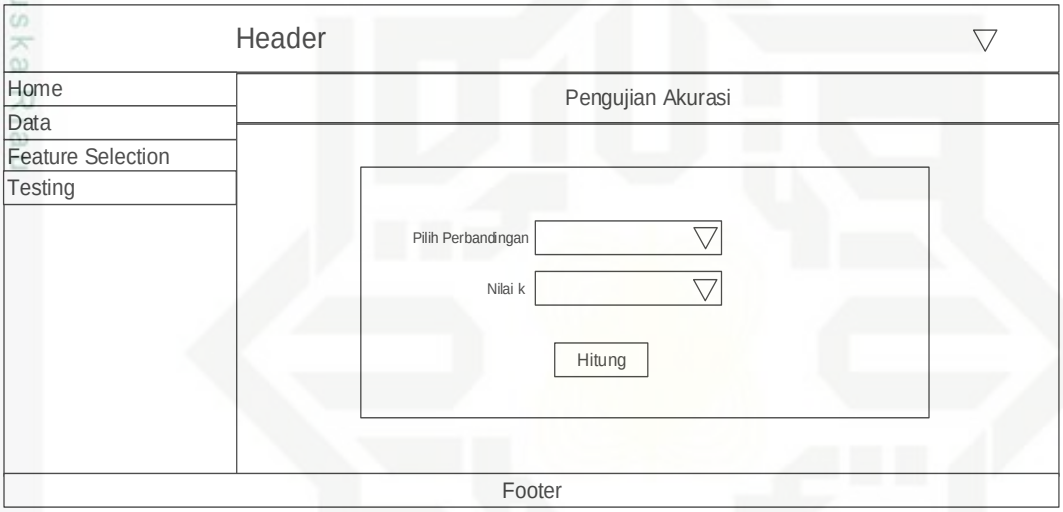


Gambar 4.9 Perancangan Halaman *Feature Selection*

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.

4.5.4 Perancangan Halaman Pengujian

Halaman pengujian ini merupakan halaman yang mempunyai *submenu* dan dapat diakses oleh *user*. Berfungsi untuk melakukan pengujian terhadap kinerja sistem dalam mengklasifikasi serangan pada dataset KDD CUP 99 dengan menggunakan metode *Modified K-Nearest Neighbor* dan *Confusion Matrix* sebagai perhitungan akurasinya. Perancangan *submenu* halaman pengujian klasifikasi dapat dilihat pada Gambar 4.14.



Gambar 4.14 Perancangan Halaman Pengujian

State Islamic University of Sultan Syarif Kasim Riau
UIN SUSKA RIAU