

BAB IV

ANALISA DAN PERANCANGAN

Pada bab ini membahas tentang analisa dan perancangan proses *feature selection* menggunakan *Symmetrical Uncertainty* dan *Gain Ratio* serta proses klasifikasi menggunakan metode *Modified K-Nearest Neighbor*. Analisa dan model klasifikasi perlu dilakukan terlebih dahulu untuk dapat menyelesaikan permasalahan yang ada. Analisa dibutuhkan untuk perancangan sistem, sementara perancangan dilakukan untuk menentukan rincian sistem yang akan dibuat.

4.1 Analisa Kebutuhan Data

Data yang digunakan pada penelitian ini adalah dataset NSL-KDD . Dataset tersebut berjumlah 125.973 data dan terdiri dari 41 atribut. Pada dataset tersebut memiliki 5 jenis kelas yang terdiri dari 1 jenis normal dan 4 jenis serangan dengan banyak data dari masing-masing jenis, yaitu *Normal* sebanyak 67.343 data, *DoS* sebanyak 45.927 data, *R2L* sebanyak 995 data, *U2R* sebanyak 52 data dan *Probe* sebanyak 11.656 data.

Jumlah kelas yang digunakan pada penelitian ini yaitu 5 kelas (*Normal*, *DoS*, *R2L*, *U2R*, *Probe*) dan data yang digunakan pada penelitian ini yaitu 1052 (tidak setara) dan 25 (setara) data, dimana nantinya akan dilihat perbandingan hasil pengujian menggunakan data yang setara dan data yang tidak setara, pada penggunaan 1052 data masing-masing kelas diambil 250 data kecuali kelas *U2R* yang hanya berjumlah 52 data, sedangkan pada data yang setara yaitu 25 data masing-masing kelas diambil 5 data dan terdapat kesetaraan setiap kelas. Data 1052 (tidak setara) dapat dilihat pada Tabel 4.1 di bawah ini (selengkapnya dapat dilihat pada Lampiran A):

Tabel 4.1 Data 1052 (tidak setara)

No	F1	F2	F3	F4	F5	...	F41	Class
1	0	tcp	ftp_data	SF	491	...	0	normal
2	240	tcp	http	SF	328	...	0	normal

No	F1	F2	F3	F4	F5	...	F41	Class
3	0	icmp	eco_i	SF	8	...	0	normal
4	0	tcp	http	SF	198	...	0.01	normal
5	7	tcp	ftp	SF	226	...	0	normal
...
1052	1	tcp	ftp	SF	1251	...	0	U2R

Adapun data yang berjumlah 25 data (setara) dapat dilihat pada Tabel 4.2 berikut:

Tabel 4.2 Data 25 (setara)

No	F1	F2	F3	F4	F5	...	F11	Class
1	0.0000	0	1.2863	0	0	...	0	normal
2	0.0385	0	8.5929	0.0000	1	...	0	normal
3	0.0385	0	5.1871	0.0013 44	1	...	0	normal
4	0.0385	0	5.6587	0.0010 26	1	...	0	normal
5	0.1154	0	2.4730	0.0000	1	...	0	normal
...
25	0.0577	0	7.9117	0.0017 23	1	...	0	U2R

4.1.1 Variabel Inputan

Analisa variabel Inputan merupakan sebuah tahapan yang dilakukan untuk menentukan parameter inputan kedalam sebuah sistem dengan tujuan untuk mendapatkan pemahaman sistem secara keseluruhan, sehingga sistem dapat berjalan sesuai dengan kebutuhan *user*. Data atau variabel inputan yang digunakan untuk proses analisa ini dapat dilihat pada Tabel 4.3.

Tabel 4.3 Keterangan Kriteria Inputan

No	Kode Atribut	Nama Fitur
1	F1	<i>Duration</i>
2	F2	<i>protocol_type</i>
3	F3	<i>Service</i>
4	F4	<i>Flag</i>
5	F5	<i>src_bytes</i>
6	F6	<i>dst_bytes</i>

No	Kode Atribut	Nama Fitur
7	F7	Land
8	F8	wrong_fragment
9	F9	Urgent
10	F10	Hot
11	F11	num_failed_logins
12	F12	logged_in
13	F13	num_compromised
14	F14	root_shell
15	F15	su_attempted
16	F16	num_root
17	F17	num_file_creations
18	F18	num_shells
19	F19	num_access_files
20	F20	num_outbound_cmds
21	F21	is_host_login
22	F22	is_guest_login
23	F23	Count
24	F24	srv_count
25	F25	serror_rate
26	F26	srv_serror_rate
27	F27	rerror_rate
28	F28	srv_rerror_rate
29	F29	same_srv_rate
30	F30	diff_srv_rate
31	F31	srv_diff_host_rate
32	F32	dst_host_count
33	F33	dst_host_srv_count
34	F34	dst_host_same_srv_rate
35	F35	dst_host_diff_srv_rate
36	F36	dst_host_same_src_port_rate
37	F37	dst_host_srv_diff_host_rate
38	F38	dst_host_serror_rate
39	F39	dst_host_srv_serror_rate
40	F40	dst_host_rerror_rate
41	F41	dst_host_srv_rerror_rate

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 Discovery in Database (KDD)

Tahapan Knowledge Discovery in Database pada penelitian ini terdiri dari:

4.2.2 Data Selection

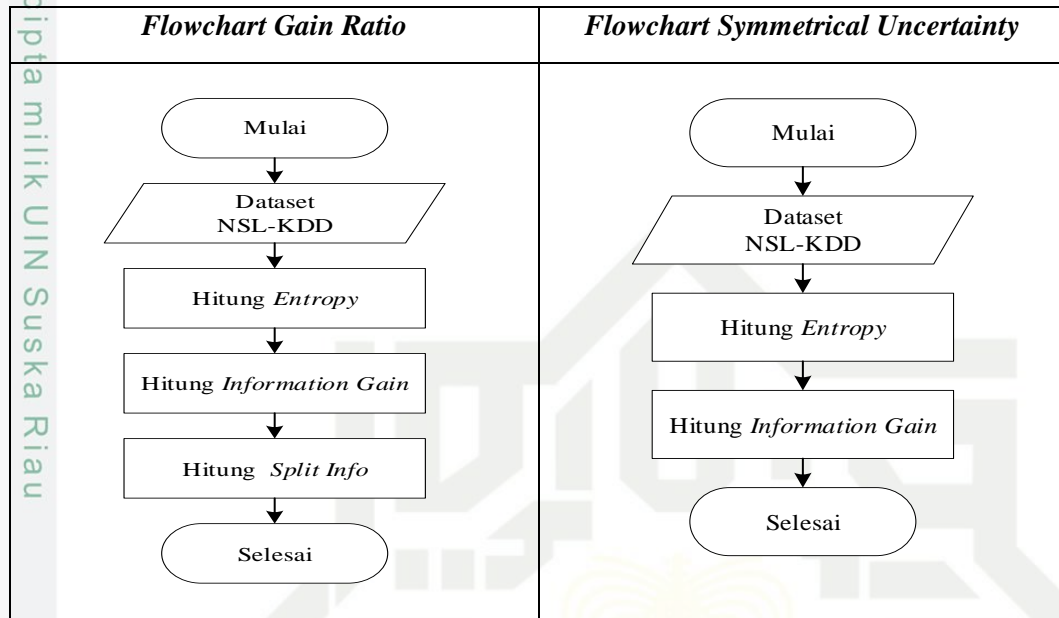
Tahapan data selection ini yaitu dilakukannya proses *feature selection* untuk mengurangi jumlah fitur yang terdapat pada dataset. *Feature selection* yang digunakan yaitu *symmetrical uncertainty* (Persamaan 2.1) dan *gain ratio* (Persamaan 2.5). Berikut proses untuk memperoleh nilai bobot pada masing-

Hak Cipta Dilindungi Undang-Undang

1. Diarangi 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.

masing atribut menggunakan *symmetrical uncertainty* dan *gain ratio* yang dapat dilihat pada Tabel 4.4 di bawah ini:

Tabel 4.4 Flowchart Symmetrical Uncertainty dan Gain Ratio



1. Menghitung Nilai Entropy Data (ID)

Perhitungan *entropy* data menggunakan (Persamaan 2.3), berikut langkah-langkah perhitungan *entropy* :

Keterangan:

Jumlah data perkelas *normal, DoS, Probe, R2L*= **250**

Jumlah data kelas *U2R*= **52**

Jumlah keseluruhan data= **1052**

$$\begin{aligned}
 Entropy(D) = & \left(\left(-\frac{250}{1052} * \log_2 \left(\frac{250}{1052} \right) \right) + \left(-\frac{250}{1052} * \log_2 \left(\frac{250}{1052} \right) \right) \right. \\
 & + \left(-\frac{250}{1052} * \log_2 \left(\frac{250}{1052} \right) \right) + \left(-\frac{250}{1052} * \log_2 \left(\frac{250}{1052} \right) \right) \\
 & \left. + \left(-\frac{52}{1052} * \log_2 \left(\frac{52}{1052} \right) \right) \right)
 \end{aligned}$$



Hak Cipta Dilindungi Undang-Undang

1. Diarangi 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.

$$Entropy = 0,4927 + 0,4927 + 0,4927 + 0,4927 + 0,2144$$

$$Entropy(D) = \mathbf{2,1851}$$

2. Menghitung Nilai Entropy Atribut (IA)

Perhitungan *entropy* atribut menggunakan (Persamaan 2.4), berikut langkah-langkah perhitungan *entropy* atribut:

Entropy (duration)

Keterangan:

Jumlah partisi “0” pada atribut “F1 (*duration*)”= **857**

Jumlah keseluruhan data= **1052**

Nilai *Entropy* dalam partisi “0” = **2.0733**

*Nilai setiap perhitungan menyesuaikan data masing-masing seperti keterangan di atas.

$$Entropy(0) = \left(\frac{857}{1052} * 2,0733 \right) = \mathbf{1.6890}$$

$$Entropy(5607) = \left(\frac{1}{1052} * 0 \right) = 0$$

$$Entropy(507) = \left(\frac{1}{1052} * 0 \right) = 0$$

...

$$Entropy(290) = \left(\frac{1}{1052} * 0 \right) = 0$$

Kemudian *entropy (0)* hingga *entropy (290)* dijumlahkan.

$$Entropy(A) = 1,6889 + 0 + 0 + \dots + 0 = \mathbf{1,7811}$$

Entropy (protocol_type)

Keterangan:

Jumlah partisi “*tcp*” pada atribut (*protocol_type*)”= **869**

Jumlah keseluruhan data= **1052**

Nilai *Entropy* dalam partisi “*tcp*” = **2.1655**

*Nilai setiap perhitungan menyesuaikan data masing-masing seperti keterangan di atas.



$$Entropy(tcp) = \left(\frac{869}{1052} * 2,1655\right) = 1,7888$$

$$Entropy(udp) = \left(\frac{72}{1052} * 1,5328\right) = 0,1049$$

$$Entropy(icmp) = \left(\frac{111}{1052} * (0,7026)\right) = 0,0741$$

Kemudian *entropy* (1) hingga *entropy* (3) dijumlahkan.

$$Entropy(A) = 1,7888 + 0,1049 + 0,0741 = \mathbf{1,9678}$$

Entropy service

Keterangan:

Jumlah partisi “*ftp_data*” pada atribut (*service*)= **192**

Jumlah keseluruhan data= **1052**

Nilai *Entropy* dalam partisi “*ftp_data*” = **1,0227**

*Nilai setiap perhitungan menyesuaikan data masing-masing seperti keterangan di atas.

$$Entropy(ftp_data) = \left(\frac{192}{1052} * 1,0227\right) = 0,1866$$

$$Entropy(other) = \left(\frac{38}{1052} * 1,3528\right) = 0,0489$$

$$Entropy(http) = \left(\frac{162}{1052} * 0,4045\right) = 0,0623$$

$$Entropy(pm_dump) = \left(\frac{1}{1052} * 0\right) = 0$$

Kemudian *entropy* (1) hingga *entropy* (53) dijumlahkan.

$$Entropy(A) = 0,1866 + 0,0489 + 0,0623 + \dots + 0 = \mathbf{0,7799}$$

Dan *entropy* selanjutnya hingga **Entropy *dst_ost_srv_rerror***.

3. Menghitung Nilai Information Gain (IG)

Perhitungan *information gain* (IG) menggunakan (Persamaan 2.2), berikut langkah-langkah perhitungan *information gain* :

$$IG(duration) = 2,1851 - 1,7811 = \mathbf{0,4041}$$

$$IG(protocol_type) = 2,1851 - 1,9678 = \mathbf{0,2173}$$

Hak Cipta Diindungi 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.

$$IG(service) = 2,1851 - 0,7799 = \mathbf{1,4053}$$

Dan seterusnya.

4. Menghitung Nilai *Split Info*

Perhitungan *split info* menggunakan (Persamaan 2.6), berikut langkah-langkah perhitungan *split info* :

Split Info duration

$$Split\ info(0) = \left(-\frac{857}{1052} * \log_2 \left(\frac{857}{1052} \right) \right) = 0,2409$$

$$Split\ info(5607) = \left(-\frac{1}{1052} * \log_2 \left(\frac{1}{1052} \right) \right) = 0,0095$$

$$Split\ info(507) = \left(-\frac{1}{1052} * \log_2 \left(\frac{1}{1052} \right) \right) = 0,0095$$

$$Split\ info(290) = \left(-\frac{1}{1052} * \log_2 \left(\frac{1}{1052} \right) \right) = 0,0095$$

Kemudian *split info* (0) hingga *split info* (290) dijumlahkan.

$$Split\ Info(duration) = \mathbf{1,7182}$$

Split Info protocol_type

$$Split\ info(tcp) = \left(-\frac{869}{1052} * \log_2 \left(\frac{869}{1052} \right) \right) = 0,2277$$

$$Split\ info(udp) = \left(-\frac{72}{1052} * \log_2 \left(\frac{72}{1052} \right) \right) = 0,2648$$

$$Split\ info(icmp) = \left(-\frac{111}{1052} * \log_2 \left(\frac{111}{1052} \right) \right) = 0,3423$$

Kemudian *split info* (1) hingga *split info* (3) dijumlahkan.

$$Split\ Info(protocol_type) = \mathbf{0,8349}$$

Split Info service

$$Split\ info(ftp_data) = \left(-\frac{194}{1052} * \log_2 \left(\frac{194}{1052} \right) \right) = 0,4479$$

$$Split\ info(other) = \left(-\frac{38}{1052} * \log_2 \left(\frac{38}{1052} \right) \right) = 0,1731$$

$$Split\ info(http) = \left(-\frac{162}{1052} * \log_2 \left(\frac{162}{1052} \right) \right) = 0,4156$$

② *Split info*(pm_dump) = $\left(-\frac{1}{1052} * \log_2\left(\frac{1}{1052}\right)\right) = 0,0095$

Kemudian *split info* (0) hingga *split info* (53) dijumlahkan.

Split Info(service) = **3,6932**

Dan *split info* selanjutnya sampai *Entropy dst_ost_srv_error*.

5. Menghitung Gain Ratio (GR)

Perhitungan *gain ratio* menggunakan Persamaan 2.1, berikut langkah-langkah perhitungan *gain ratio* :

$GR(duration) = \frac{0,4041}{1,7182} = \mathbf{0,2352}$

$GR(protocol_type) = \frac{0,2173}{0,8349} = \mathbf{0,2603}$

$GR(service) = \frac{1,4053}{3,6932} = \mathbf{0,3805}$

Dan seterusnya hingga *Fitur dst_ost_srv_error*.

6. Menghitung Symmetrical Uncertainty (SU)

Perhitungan *symmetrical uncertainty* menggunakan (Persamaan 2.6), berikut langkah-langkah perhitungan *symmetrical uncertainty* :

$SU(duration) = 2 * \left(\frac{0,4041}{2,1851+1,7811}\right) = \mathbf{0,2037}$

$SU(protocol_type) = 2 * \left(\frac{0,2173}{2,1851+1,9678}\right) = \mathbf{0,1047}$

$SU(service) = 2 * \left(\frac{1,4053}{2,1851+0,7799}\right) = \mathbf{0,9479}$

Dan seterusnya hingga *Fitur dst_ost_srv_error*.

Hasil perhitungan *feature selection symmetrical uncertainty* dan *gain ratio* dapat dilihat pada Tabel 4.5 di bawah ini :



Tabel 4.5 Hasil Perhitungan *Feature Selection*

No	Feature	Partisi	Jmlh kasus	Normal	Dos	Probe	R2L	U2R	Entropy Data (ID)	Entropy Atribut (IA)	IG	SPLIT	GR	SU
	duration		1052	250	250	250	250	52						
(F1)	0	857	226	250	208	158	15	2.0733	1.6890		0.2409			
	5607	1	1	0	0	0	0	0.0000	0.0000		0.0095			
	507	1	1	0	0	0	0	0.0000	0.0000		0.0095			

	290	1	0	0	0	0	1	0.0000	0.0000		0.0095			
(F2)	protocol_type		1052	250	250	250	250	52	2.1851	1.9678	0.2173	0.8349	0.2603	0.1047
	tcp	869	204	233	133	250	49	2.1655	1.7888		0.2277			
	udp	72	40	8	21	0	3	1.5328	0.1049		0.2648			
	icmp	111	6	9	96	0	0	0.7026	0.0741		0.3423			
(F3)	service		1052	250	250	250	250	52	2.1851	0.7799	1.4053	3.6932	0.3805	0.9479
	ftp_data	192	17	7	1	155	12	1.0227	0.1866		0.4479			
	other	38	10	0	24	1	3	1.3528	0.0489		0.1731			
	http	162	151	9	0	2	0	0.4045	0.0623		0.4156			

	pm_dump	1	0	0	1	0	0	0.0000	0.0000		0.0095			

(F41)	dst_host_srv_rerror_rate		1052	250	250	250	250	52	2.1851	1.8920	0.2931	1.0710	0.2736	0.1438
	0	847	227	207	130	238	45	2.1605	1.7395		0.2518			
	0.01	8	6	2	0	0	0	0.8113	0.0062		0.0535			

	0.14	1	0	0	0	0	1	0.0000	0.0000		0.0095			

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa

d. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

7. Menentukan Pilihan Atribut/Fitur

Setelah diperoleh nilai bobot pada masing-masing fitur, selanjutnya dipilih atribut yang akan digunakan untuk klasifikasi MK-NN. Pemilihan fitur dilakukan dengan cara perangkingan nilai dari yang terbesar hingga terkecil berdasarkan penelitian yang telah dilakukan sebelumnya oleh (Gark and Kumar, 2014). Hasil perangkingan dapat dilihat pada Tabel 4.6 di bawah ini:

Tabel 4.6 Perangkingan SU dan GR

RANKING SU			RANKING GR		
No	Fitur ke-	Nilai GR	No	Fitur ke-	Nilai SU
1	5	1.2172	1	12	1.2960
2	3	0.9479	2	14	0.6060
3	33	0.6383	3	26	0.5262
4	23	0.5922	4	4	0.4616
5	6	0.5896	5	25	0.4364
6	35	0.4939	6	39	0.4198
7	4	0.4149	7	17	0.4169
8	30	0.4061	8	22	0.4141
9	34	0.4035	9	13	0.3885
10	32	0.4035	10	30	0.3874
11	37	0.3782	11	3	0.3805
12	36	0.3610	12	9	0.3790
13	29	0.3509	13	16	0.3664
14	24	0.3256	14	18	0.3543
15	12	0.3213	15	10	0.3376
16	38	0.2948	16	38	0.3260
17	39	0.2762	17	5	0.3231
18	25	0.2718	18	29	0.2820
19	26	0.2378	19	6	0.2812
20	1	0.2037	20	41	0.2736
21	40	0.1840	21	2	0.2603
22	41	0.1438	22	37	0.2511
23	10	0.1427	23	35	0.2478
24	31	0.1420	24	31	0.2451
25	2	0.1047	25	8	0.2376

Hak Cipta Diindungi 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 mengummumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

RANKING SU		
No	Fitur ke-	Nilai GR
26	27	0.0956
27	28	0.0777
28	22	0.0728
29	14	0.0518
30	13	0.0491
31	17	0.0450
32	16	0.0224
33	11	0.0107
34	18	0.0091
35	8	0.0082
36	19	0.0035
37	9	0.0019
38	15	0.0009
39	7	0.0000
40	20	0.0000
41	21	0.0000

RANKING GR		
No	Fitur ke-	Nilai SU
26	1	0.2352
27	11	0.2309
28	36	0.2212
29	23	0.2154
30	28	0.2121
31	34	0.1974
32	27	0.1944
33	32	0.1843
34	40	0.1823
35	33	0.1816
36	15	0.1808
37	24	0.1580
38	19	0.1243
39	7	0.0000
40	20	0.0000
41	21	0.0000

Setelah dirangkingkan, maka pemilihan fitur dilakukan dengan menetapkan nilai 20 tertinggi pada nilai *gain ratio* dan *symmetrical uncertainty* berdasarkan penelitian yang telah dilakukan sebelumnya oleh (Garg and Kumar, 2014) dengan judul *Combinational Feature Selection Approach for Network Intrusion Detection System* yang telah dijelaskan sebelumnya pada Bab 3.3.4.1. Hasil *feature selection* dengan pengambilan 20 nilai tertinggi dapat dilihat pada Tabel 4.7 di bawah ini :

Tabel 4.7 20 Nilai Tertinggi SU dan GR

RANKING SU		
No	Fitur ke-	Nilai GR
1	5	1.2172
2	3	0.9479
3	33	0.6383
4	23	0.5922
5	6	0.5896
6	35	0.4939

RANKING GR		
No	Fitur ke-	Nilai SU
1	12	1.2960
2	14	0.6060
3	26	0.5262
4	4	0.4616
5	25	0.4364
6	39	0.4198

Hak Cipta Dilindungi Undang-Undang

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

RANKING SU		
No	Fitur ke-	Nilai GR
7	4	0.4149
8	30	0.4061
9	34	0.4035
10	32	0.4035
11	37	0.3782
12	36	0.3610
13	29	0.3509
14	24	0.3256
15	12	0.3213
16	38	0.2948
17	39	0.2762
18	25	0.2718
19	26	0.2378
20	1	0.2037

RANKING GR		
No	Fitur ke-	Nilai SU
7	17	0.4169
8	22	0.4141
9	13	0.3885
10	30	0.3874
11	3	0.3805
12	9	0.3790
13	16	0.3664
14	18	0.3543
15	10	0.3376
16	38	0.3260
17	5	0.3231
18	29	0.2820
19	6	0.2812
20	41	0.2736

Setelah mendapatkan 20 nilai tertinggi, maka selanjutnya dilakukan proses *boolean AND operator* dengan cara mencari irisan atau keterkaitan fitur dari kedua *feature selection* hingga memperoleh fitur untuk di klasifikasi, untuk lebih jelasnya akan ditampilkan pada Tabel 4.8 dan Tabel 4.9 berikut:

Tabel 4.8 Boolean And Operator

No	SU	GR
1	F5	F12
2	F3	F14
3	F33	F26
4	F23	F4
5	F6	F25
6	F35	F39
7	F4	F17
8	F30	F22
9	F34	F13
10	F32	F30
11	F37	F3
12	F36	F9

Tabel 4.9 Hasil Feature Selection

No	SU + GR
1	F3
2	F4
3	F5
4	F6
5	F12
6	F25
7	F26
8	F29
9	F30
10	F38
11	F39

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.

13	F29	F16
14	F24	F18
15	F12	F10
16	F38	F38
17	F39	F5
18	F25	F29
19	F26	F6
20	F1	F41

Setelah dilakukan tahap *boolean AND operator*, maka diperoleh hasil dari kedua metode *feature selection* yang digunakan yaitu *symmetrical uncertainty* dan *gain ratio* sebanyak 11 fitur, pada Tabel 4.7 dan Tabel 4.8 tersebut, yang diberi warna kuning merupakan fitur yang saling berkaitan dan yang akan digunakan untuk klasifikasi.

4.2.3 Data Transformation

Tahapan transformasi ini dilakukannya perubahan nilai atribut yang bertipe nominal menjadi numerik. Atribut yang bertipe nominal yaitu *protocol_type*, *service*, *flag* dan *class*. Berikut ini perubahan nilai atribut yang dapat dilihat pada Tabel 4.10 di bawah ini :

Tabel 4.10 Transformasi Data

Atribut	Nilai Atribut	Transformasi
Protocol_type (F2)	tcp	1
	udp	2
	icmp	3
Service (F3)	ftp_data	1
	other	2
	http	3

	pm_dump	53
Flag (F4)	SF	1
	REJ	2
	RSTO	3

	S1	9
Class	Normal	1

Hak Cipta Diindungi Undang-Undang

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

Atribut	Nilai Atribut	Transformasi
(F42)	Dos	2
	Probe	3
	R2L	4
	U2R	5

4.2.4 Normalisasi

Tahapan normalisasi dilakukannya proses merubah data numerik yang telah ditransformasi menjadi *range* data berkisar 0 – 1, dengan tujuan agar sebaran data tidak jauh. Normalisasi dilakukan menggunakan Persamaan (2.15). Berikut adalah proses perhitungan normalisasi :

Pada kolom *duration*, dari 1052 data, nilai minimum = 0 dan nilai maximum = 41802

$$\text{Data ke-1} \quad X^* = \frac{0-0}{41802-0} = 0$$

....

$$\text{Data ke-1052} \quad X^* = \frac{1-0}{41802-0} = 2.39E-05$$

Pada kolo *protocol_type*, dari 1052 data, nilai minimum = 1 dan nilai maximum =3

$$\text{Data ke-1} \quad X^* = \frac{1-1}{3-1} = 0$$

....

$$\text{Data ke-1052} \quad X^* = \frac{1-1}{3-1} = 0$$

Dan seterusnya hingga Fitur *dst_host_srv_error_rate*.

4.3 Klasifikasi *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 jelasnya mengenai cara kerja algoritma MK-NN akan dijelaskan di bawah ini.

1. Pembagian data latih dan data uji

Untuk melakukan klasifikasi, dibutuhkan data latih sebagai data pembelajaran untuk menentukan kelas pada data uji. Data latih yang digunakan adalah data latih yang telah dinormaslisasi. Pada pembagian data untuk perhitungan manual menggunakan data yang berjumlah 1052 data sedangkan data yang

Hak Cipta Diindungi Undang-Undang

1. Diarangi 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.

berjumlah 25 tidak dilakukan perhitungan manual dikarenakan tidak terdapat perbedaan pada setiap proses perhitungan manual. Adapun data yang akan digunakan sebagai data latih dalam contoh perhitungan manual diambil 90% dari masing-masing kelas sehingga data latih untuk penelitian ini adalah 947 (90%:10%). Berikut data latih dapat dilihat pada Tabel 4.11 :

Tabel 4.11 Data Latih

No	X1	X2	X3	X4	...	X11	Class
1	0	0	1.28631	0	...	0	1
2	0.03846	0	8.59293	0.00005	...	0	1
3	0.03846	0	5.18719	0.00134	...	0	1
4	0.03846	0	5.65875	0.00102	...	0	1
5	0.11538	0	2.47308	0.00006	...	0	1
6	0.07692	0	8.64532	0	...	0	1
7	0.03846	0	7.75459	0.00110	...	0	1
8	0.03846	0	4.61084	0.00008	...	0	1
9	0.03846	0	6.10412	0.00047	...	0	1
...
947	0.05769	0.25	3.94541	0.00036	...	0	5

Data uji merupakan data yang akan uji pada sistem untuk kebutuhan klasifikasi data serangan terhadap data latih. Pengujian dilakukan untuk menentukan tingkat akurasi proses klasifikasi, data uji yang digunakan pada penelitian ini berjumlah 10% dari masing-masing kelas sehingga jumlah data uji sebanyak 105 data dapat dilihat pada Tabel 4.12 berikut:

Tabel 4.12 Data Uji

No	X1	X2	X3	X4	X11	Class
1	0.11538	0	2.49928	0.00006	0	1
2	0.09615	0	2.75078	0.00002	0	1
3	0.03846	0	7.49261	0.00027	0	1
4	0.03846	0	6.78527	0.00014	0	1
5	0.09615	0	2.75078	0.00002	0	1
....
105	0.05769	0	7.25683	0.00013	0	5

Tahap selanjutnya dilakukan proses perhitungan dengan menggunakan MK-NN. Proses MK-NN dibagi menjadi empat tahap perhitungan, yaitu jarak *Euclidean* antar data latih, jarak *Euclidean* data latih dengan data uji, *Validitas* 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 (2.6), berikut langkah – langkah perhitungan jarak *euclidean* data latih :

a. Untuk jarak antara data latih 1 dan 2

$$d(x, y)(\text{data latih } x, \text{ data latih } y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$d(1,2) = \sqrt{\begin{matrix} (0 - 0.03846)^2 + (0 - 0)^2 + (1.28631 - 8.59293)^2 \\ + \\ (0 - 0.00005)^2 + (0 - 1)^2 + (0 - 0)^2 + (0 - 0)^2 + (1 - 1)^2 \\ + (0 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 \end{matrix}}$$

$$d(1,2) = 1.00074$$

b. Untuk Jarak Antara Data Latih 1 dan 3.

$$d(x, y)(\text{data latih } x, \text{ data latih } y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$d(1,3) = \sqrt{\begin{matrix} (0 - 0.03846)^2 + (0 - 0)^2 + (1.28631 - 5.18719)^2 \\ + (0 - 0.00134)^2 + \\ (0 - 1)^2 + (0 - 0)^2 + (0 - 0)^2 + (1 - 1)^2 + (0 - 0)^2 \\ + (0 - 0)^2 + (0 - 0)^2 \end{matrix}}$$

$$d(1,3) = 1.00074$$

c. Untuk Jarak Antara Data Latih 2 dan 3.

$$d(x, y)(\text{data latih } x, \text{ data latih } y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$d(2,3) = \sqrt{\begin{matrix} (0.03846 - 0.03846)^2 + (0 - 0)^2 \\ + (8.59293 - 5.18719)^2 \\ + (0.00005 - 0.00134)^2 + \\ (1 - 1)^2 + (0 - 0)^2 + (0 - 0)^2 + (1 - 1)^2 + (0 - 0)^2 \\ + (0 - 0)^2 + (0 - 0)^2 \end{matrix}}$$

$$d(2,3) = 0.00129$$

Perhitungan diatas dilakukan pada semua data latih untuk memperoleh jarak *euclidean* secara keseluruhan. Perhitungan dilakukan dengan cara dan persamaan

Hak Cipta Diindungi 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 mengummumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

yang sama. Nilai dari perhitungan jarak *euclidean* data latih dapat dilihat pada Tabel 4.13. Diketahui P adalah sampel data.

Tabel 4.13 Nilai Jarak Antar Data Latih

P	D1	D2	D3	D4	D5	D947
D1	0	1.00074	1.00074	1.00074	1.006635	1.74058
D2	1.00074	0	0.00129	0.00097	0.076923	1.42361
D3	1.00074	0.00129	0	0.00032	0.076934	1.42361
D4	1.00074	0.00097	0.00032	0	0.076929	1.42361
D5	1.00663	0.07692	0.07693	0.07693	0	1.42465
....
D947	1.74058	1.42361	1.42361	1.42361	1.42465	0

3. Menghitung Nilai Validitas Data Latih

Menghitung validitas pada setiap data latih menggunakan Persamaan (2.7) pemberian nilai 1 atau 0 untuk setiap ketetanggaan menggunakan Persamaan (2.8). Sebelum melakukan perhitungan, tentukan nilai k atau jarak ketetanggaannya terlebih dahulu. Pada contoh perhitungan ini ditentukan nilai $k = 3$ berikut adalah langkah-langkah untuk mencari nilai validitasnya:

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

$$Validitas (data_1) = \frac{1}{3} * (1 + 1 + 1) = 1$$

$$Validitas (data_2) = \frac{1}{3} * (1 + 1 + 1) = 1$$

.....

$$Validitas (data_{947}) = \frac{1}{3} * (0 + 0 + 1) = 0.33333$$

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 4.14 (selengkapnya dapat dilihat pada lampiran).

Tabel 4.14 Validitas Data Latih

Data Ke-	Validitas
1	1
2	1
3	1
4	1
5	1
....
947	0.33333

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 4.8 dan Tabel 4.9, berikut langkah perhitungan jarak *euclidean* data uji dengan data latih :

$$d(\text{datauji}x, \text{datalatih}x) = \sqrt{\sum_{i=1}^n (x_{2i} - x_{1i})^2}$$

$$d(\text{datauji1}, \text{datalatih1})$$

$$= \sqrt{\begin{aligned} &(0.11538 - 0)^2 + (0 - 0)^2 + (2.49928 - 1.28631)^2 \\ &+ (0.00006 - 0)^2 + (1 - 0)^2 + (0 - 0)^2 \\ &+ (0 - 0)^2 \\ &+ (0.5 - 1)^2 + (1 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 \end{aligned}}$$

$$= \mathbf{1.50443}$$

$$d(\text{datauji2}, \text{datalatih1})$$

$$= \sqrt{\begin{aligned} &(0.09615 - 0)^2 + (0 - 0)^2 + (2.75078 - 1.28631)^2 \\ &+ (0.00002 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 \\ &+ (0 - 0)^2 \\ &+ (1 - 1)^2 + (0 - 0)^2 + (0 - 0)^2 + (0 - 0)^2 \end{aligned}}$$

$$= \mathbf{0.09615}$$

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 4.15.

Tabel 4.15 Nilai Jarak *Euclidean* Data Uji dengan Data Latih

D	DU 1	DU 2	DU 3	...	DU 105
1	1.50443	0.09615	1.00074	...	0.05769
2	1.12068	1.00166	0.00022	...	1.00018
3	1.12068	1.00166	0.00107	...	1.00018
4	1.12068	1.00166	0.00075	...	1.00019
5	1.11803	1.00018	0.07692	...	1.00166
...
947	1.47297	1.74005	1.42361	...	1.73963

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 (2.9), mengacu pada Tabel 4.11 dan Tabel 4.12, berikut langkah – langkah perhitungan *weight voting* antara data uji dengan data latih :

$$W(x) = Validitas(x) * \frac{1}{d_x + 0,5}$$

$$W(\text{datauji1}) = 1 * \frac{1}{1.50443 + 0.5} = \mathbf{0.0.49889}$$

$$W(\text{datauji2}) = 1 * \frac{1}{0.09615 + 0.5} = \mathbf{0.61702}$$

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 4.16.

Tabel 4.16 Nilai *Weight Voting*

No	WV1	WV2	WV3	WV4	WV5	...	WV947
1	0.49889	0.61702	0.61702	0.61702	0.61803	0.16895

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 1052 data yang telah dilakukan perhitungan dapat dilihat pada Tabel 4.17

Tabel 4.17 Urutan Nilai Weight Voting Dari Yang Terbesar Hingga Terkecil

No	WV Data Uji	WV Data Uji	WV Data Uji
1	1.27625 WV:171 Class: 1	1.12994 WV:168 Class:1	1.08333 WV:837 Class: 4

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 4.18 di bawah ini :

Tabel 4.18 Kelas Asli Hasil Weight Voting

No	WV Data Uji	WV Data uji	WV Data Uji	Mayoritas
1	Normal	Normal	R2L	Normal

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 Tabel 4.19 di bawah ini :

Tabel 4.19 Hasil Klasifikasi

No	Kelas Asli	Kelas Hasil Klasifikasi	Terkenali
1	Normal	Normal	Benar

4.4 Analisa Fungsional Sistem

Analisa fungsional sistem akan menjelaskan sistem yang akan dibangun dengan *usecase diagram*, *usecase specification*, *activity diagram*, *sequence diagram*, *class diagram* dan *deployment diagram*.

4.4.1 Usecase Diagram

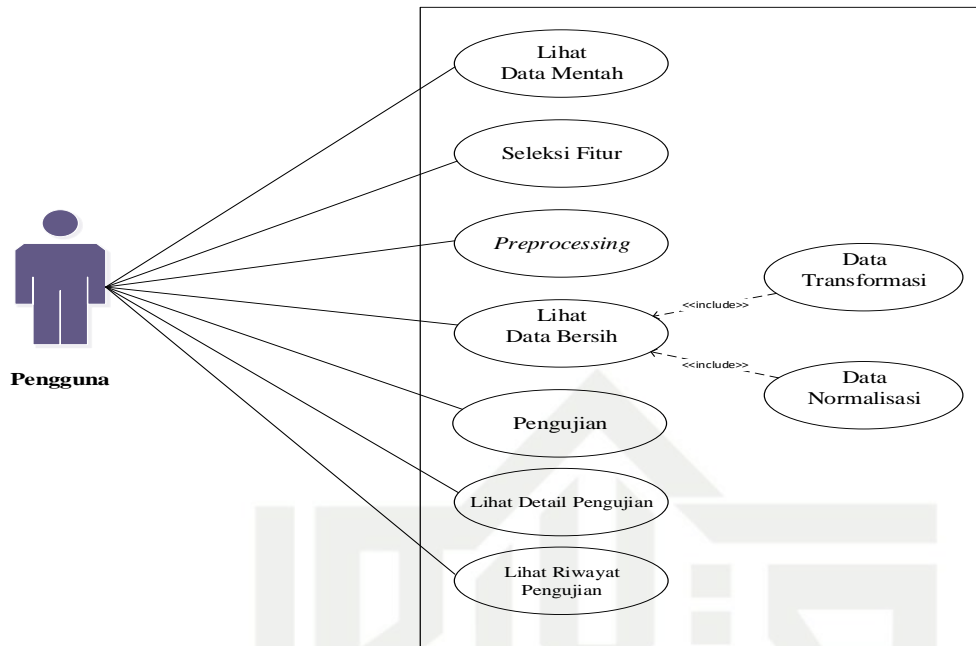
Usecase diagram akan menggambarkan aktivitas dari setiap pengguna dalam sistem yang dapat dilihat pada Gambar 4.1 berikut:

Hak Cipta Diindungi 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 4.1 Usecase Diagram

4.4.2 Usecase Specification

Usecase specification akan menjelaskan rincian dari setiap proses pada usecase diagram. Agar lebih jelas akan ditampilkan pada Tabel 4.19 di bawah ini:

1. Usecase Specification Lihat Data

Usecase specification untuk proses lihat data dapat dilihat pada Tabel 4.20 berikut:

Tabel 4.20 Usecase Specification Lihat Data

Aktor Utama	Pengguna
Kondisi Awal	Aktor ingin melihat Data Mentah
Kondisi Akhir	Aktor melihat Tabel Data Mentah dan Data Bersih
Main Succes Scenario	<ol style="list-style-type: none"> 1. Usecase dimulai ketika pengguna ingin melihat data yang digunakan untuk penelitian 2. Aktor memilih menu Data Mentah <ul style="list-style-type: none"> - Sistem menampilkan Tabel data mentah yang digunakan pada penelitian.
Alternative Scenario	

2. Usecase specification Seleksi Fitur

Usecase specification untuk proses seleksi fitur dapat dilihat pada Tabel 4.21 berikut:



Tabel 4.21 Usecase Specification Seleksi Fitur

Aktor Utama	Pengguna
Kondisi Awal	Aktor ingin melihat fitur yang diseleksi.
Kondisi Akhir	Aktor melihat hasil fitur yang terseleksi.
Main <i>Succes Scenario</i>	<ol style="list-style-type: none"> 1. Usecase dimulai ketika Aktor ingin melihat hasil fitur yang telah diseleksi. 2. Aktor memilih menu <i>Feature Selection</i> 3. Sistem menampilkan beberapa menu yaitu <i>Feature Selection</i>, SUM FS, SU&GR, Perangkingan, <i>Boolean AND</i> dan HASIL FS. <ul style="list-style-type: none"> - Jika Aktor memilih menu <i>Feature Selection</i> maka sistem menampilkan Tabel data 41 fitur serta jumlah kasus dan partisi dari setiap fitur. - Jika Aktor memilih menu SUM FS maka sistem menampilkan total partisi dari setiap fitur yang ada. - Jika Aktor memilih menu SU & GR maka sistem menampilkan nilai dari masing-masing metode <i>feature selection</i> yang digunakan. - Jika Aktor memilih menu Perangkingan, maka sistem menampilkan 41 fitur dengan urutan nilai tertinggi hingga terendah. - Jika Aktor memilih menu booleand AND, maka sistem menampilkan Tabel data 20 nilai fitur tertinggi. - Jika Aktor memilih menu HASIL FS, maka sistem menampilkan Tabel data 11 fitur yang akan digunakan untuk klasifikasi.
Alternative Scenario	

3. Usecase Specification Preprocessing

Usecase specification untuk proses *preprocessing* dapat dilihat pada Tabel

4.22 berikut:

Tabel 4.22 Usecase Specification Preprocessing

Aktor Utama	Pengguna
Kondisi Awal	Aktor ingin membersihkan data.
Kondisi Akhir	Data dibersihkan.
Main <i>Succes Scenario</i>	<ol style="list-style-type: none"> 1. Usecase dimulai ketika Aktor ingin melakukan transformasi data dan normalisasi data. 2. Aktor memilih menu <i>Preprocessing</i>. 3. Sistem menampilkan informasi bahwa data berhasil ditransformasi dan dinormalisasi.
Alternative Scenario	

4. Usecase Specification Lihat Data Bersih

Usecase specification untuk proses lihat data bersih dapat dilihat pada Tabel

4.23 berikut:

Hak Cipta Diindungi 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 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 4.23 Usecase Specification Lihat Data Bersih

Aktor Utama	Pengguna
Kondisi Awal	Aktor ingin melihat data yang telah bersih.
Kondisi Akhir	Aktor melihat hasil data transformasi dan normalisasi.
Main <i>Succes Scenario</i>	<ol style="list-style-type: none"> 1. <i>Usecase</i> dimulai ketika Aktor ingin melihat data yang telah dibersihkan. 2. Aktor memilih menu Data Bersih 3. Sistem menampilkan menu <i>dropdown</i>, yaitu Data Transformasi dan Data Normalisasi <ul style="list-style-type: none"> - Jika Aktor memilih Data Transformasi maka sistem menampilkan Tabel data yang telah ditransformasikan. - Jika Aktor memilih Data Normalisasi maka sistem menampilkan Tabel data yang telah dinormalisasikan.
<i>Alternative Scenario</i>	

5. Usecase Specification Pengujian

Usecase specification untuk proses pengujian dapat dilihat pada Tabel 4.24 berikut:

Tabel 4.24 Usecase Specification Pengujian

Aktor Utama	Pengguna
Kondisi Awal	Aktor ingin melakukan pengujian.
Kondisi Akhir	Aktor melihat hasil perhitungan akurasi
Main <i>Succes Scenario</i>	<ol style="list-style-type: none"> 1. <i>Usecase</i> dimulai ketika Aktor ingin melakukan pengujian perhitungan algoritma pada sistem. 2. Aktor memilih menu Pengujian. 3. Sistem menampilkan form untuk melakukan pengujian. 4. Aktor memilih nilai K dan memilih rasio pembagian data latih dan data uji. 5. Sistem melakukan perhitungan dan klasifikasi terhadap data uji. 6. Sistem menampilkan Tabel hasil klasifikasi dan confusion matrix serta detail pengujian.
<i>Alternative Scenario</i>	

6. Usecase Spesification Detail Pengujian

Usecase specification untuk proses detail pengujian dapat dilihat pada Tabel 4.25 berikut:

Tabel 4.25 Usecase Specification Detail Pengujian

Aktor Utama	Pengguna
Kondisi Awal	Aktor ingin melihat detail pengujian.
Kondisi Akhir	Aktor melihat hasil detail pengujian.
Main <i>Succes Scenario</i>	<ol style="list-style-type: none"> 1. <i>Usecase</i> dimulai ketika Aktor ingin melihat detail hasil pengujian. 2. Aktor memilih menu Detail Pengujian. 3. Sistem menampilkan detail hasil pengujian seperti: <ul style="list-style-type: none"> - Pembagian data latih - Pembagian data uji



© Hak Cipta miliknya UIN Suska Riau	<ul style="list-style-type: none"> - <i>Euclidean</i> Data Latih dan Latih - <i>Euclidean</i> Data Latih dan Uji - Validitas - <i>Weight Voting</i>
<i>Alternative Scenario</i>	

7. Usecase Specification Lihat Riwayat Pengujian

Usecase specification untuk proses lihat riwayat pengujian dapat dilihat pada Tabel 4.26 berikut:

Tabel 4.26 Usecase Specification Lihat Riwayat Pengujian

Aktor Utama	Pengguna
Kondisi Awal	
Kondisi Akhir	Aktor melihat riwayat hasil pengujian akurasi.
Main <i>Succes Scenario</i>	<ol style="list-style-type: none"> 1. <i>Usecase</i> dimulai ketika Aktor ingin melihat riwayat hasil pengujian akurasi pada sistem. 2. Aktor memilih menu Riwayat Pengujian. 3. Sistem menampilkan beberapa riwayat pengujian yang pernah dilakukan.
<i>Alternative Scenario</i>	

4.4.3 Activity Diagram

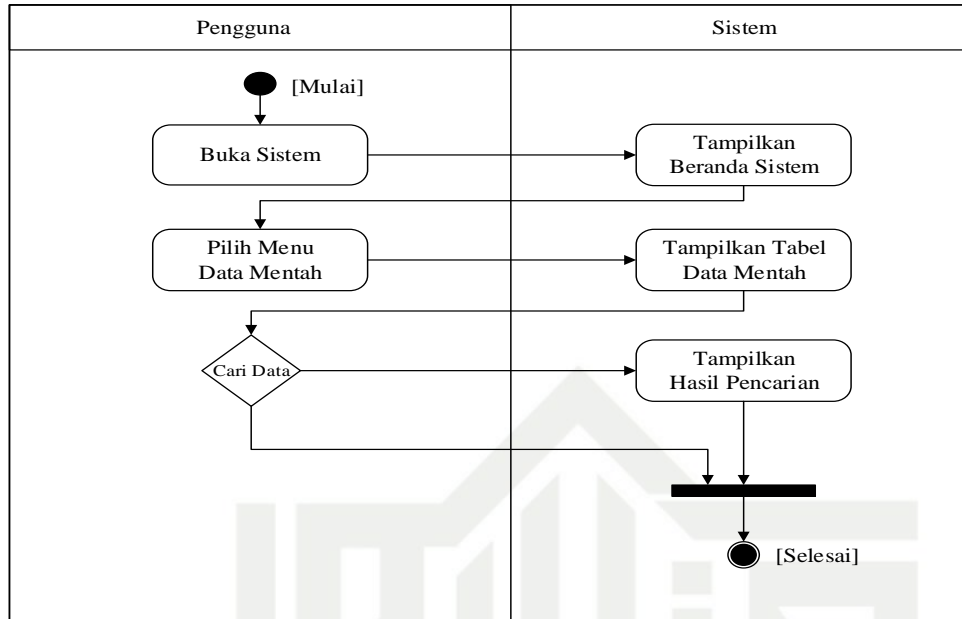
Activity diagram akan menggambarkan aliran kerja dari fungsional sistem atau proses bisnis. *Activity diagram* menggambarkan aktivitas sistem, bukan apa yang dilakukan oleh Aktor melainkan sistem. Untuk *activity diagram* pada sistem klasifikasi serangan jaringan ini dapat dilihat pada Gambar 4.2 hingga Gambar 4.9:

1. Lihat Data Mentah

Untuk *activity diagram* proses pada saat pengguna ingin mengakses halaman data mentah pada sistem klasifikasi serangan jaringan ini dapat dilihat pada Gambar 4.2 berikut:

Hak Cipta Dilindungi Undang-Undang

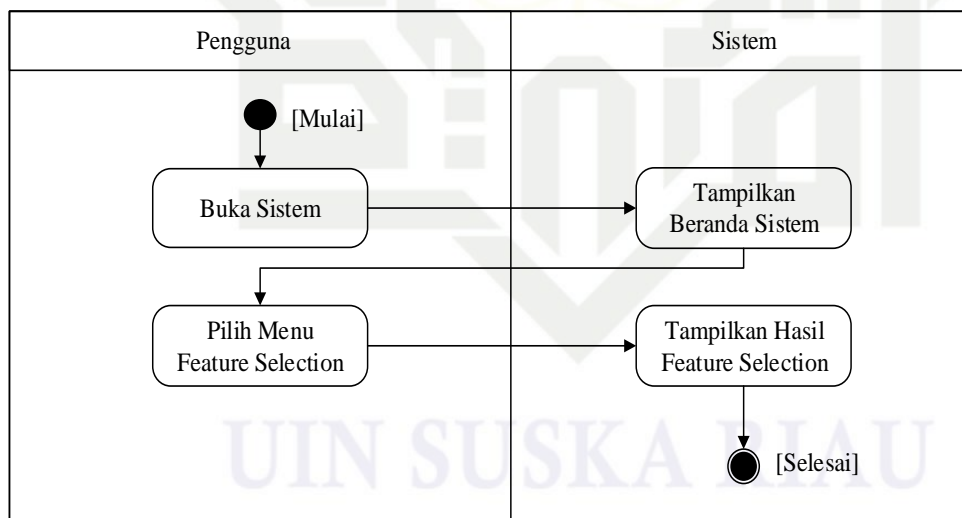
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan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



Gambar 4.2 Activity Diagram Lihat Data Mentah

2. Seleksi Fitur

Untuk *activity diagram* proses pada saat pengguna ingin melihat hasil *feature selection* pada sistem klasifikasi serangan jaringan ini dapat dilihat pada Gambar 4.3 berikut:



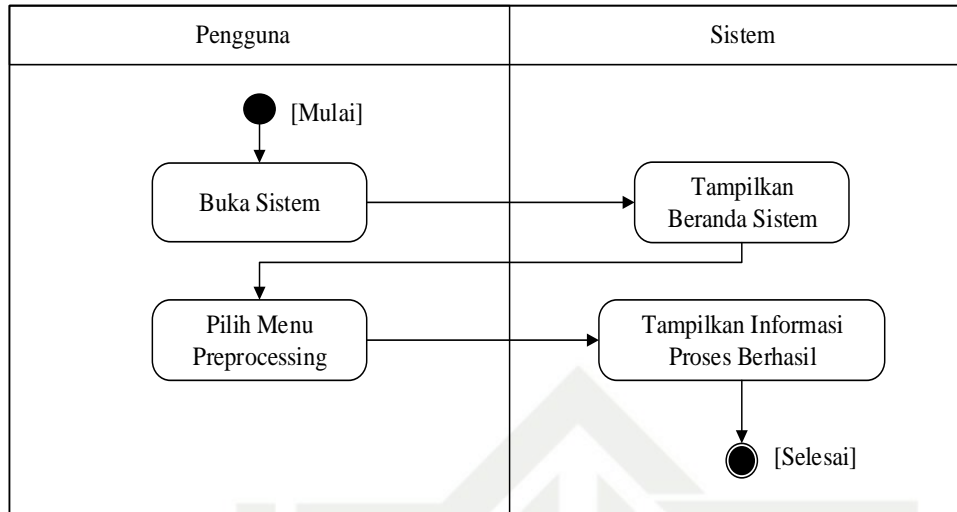
Gambar 4.3 Activity Diagram Seleksi Fitur

3. Preprocessing

Untuk *activity diagram* proses pada saat pengguna ingin melakukan tahap *preprocessing* pada sistem klasifikasi serangan jaringan ini dapat dilihat pada Gambar 4.4 berikut:

Hak Cipta Dilindungi Undang-Undang

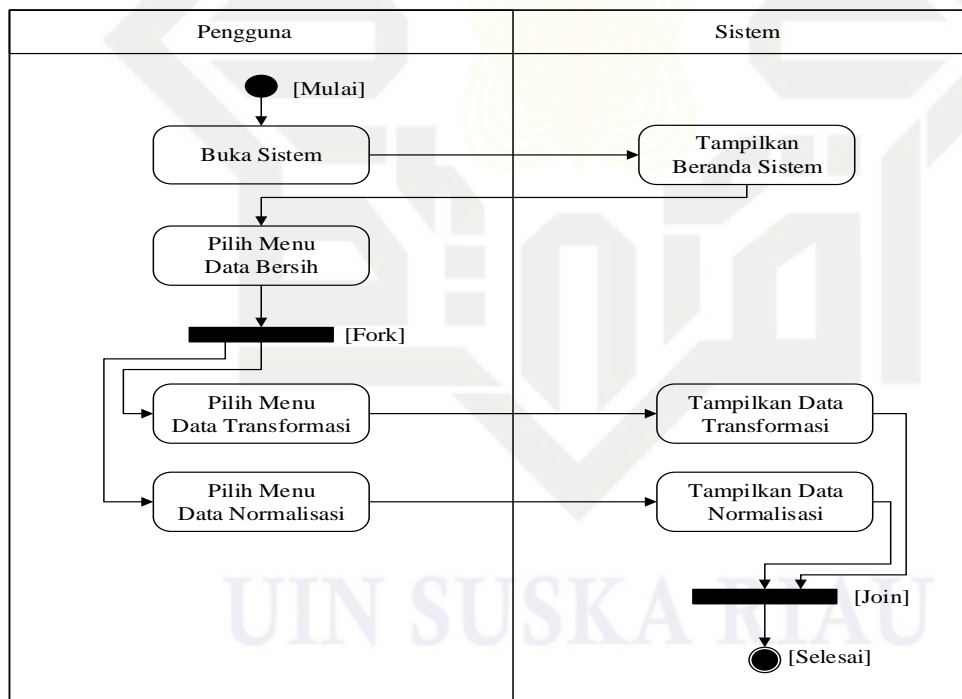
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan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



Gambar 4.4 Activity Diagram Preprocessing

4. Lihat Data Bersih

Untuk *activity diagram* proses pada saat pengguna ingin melihat data transformasi dan data normalisasi pada sistem klasifikasi serangan jaringan ini dapat dilihat pada Gambar 4.5 berikut:



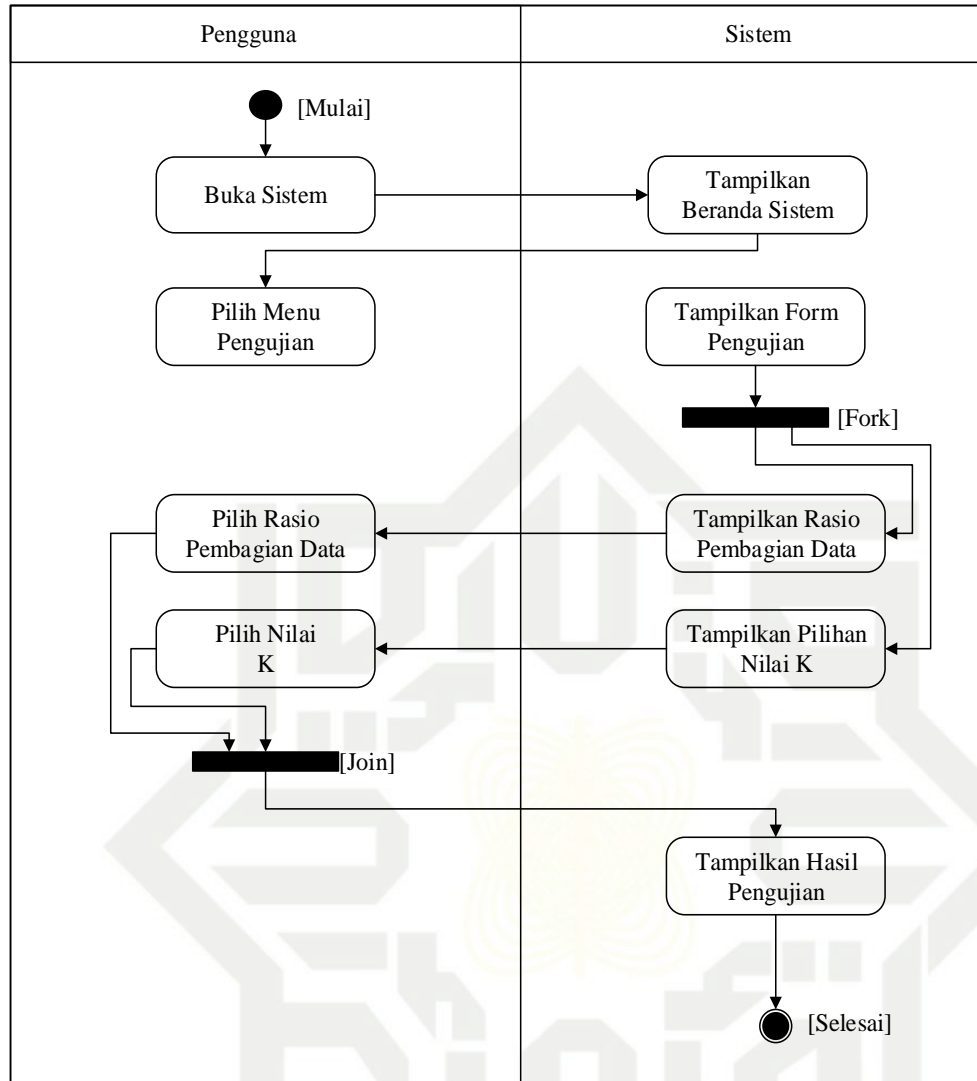
Gambar 4.5 Activity Diagram Data Bersih

5. Pengujian

Untuk *activity diagram* proses pada saat pengguna ingin melakukan pengujian pada sistem klasifikasi serangan jaringan ini dapat dilihat pada Gambar 4.6 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.



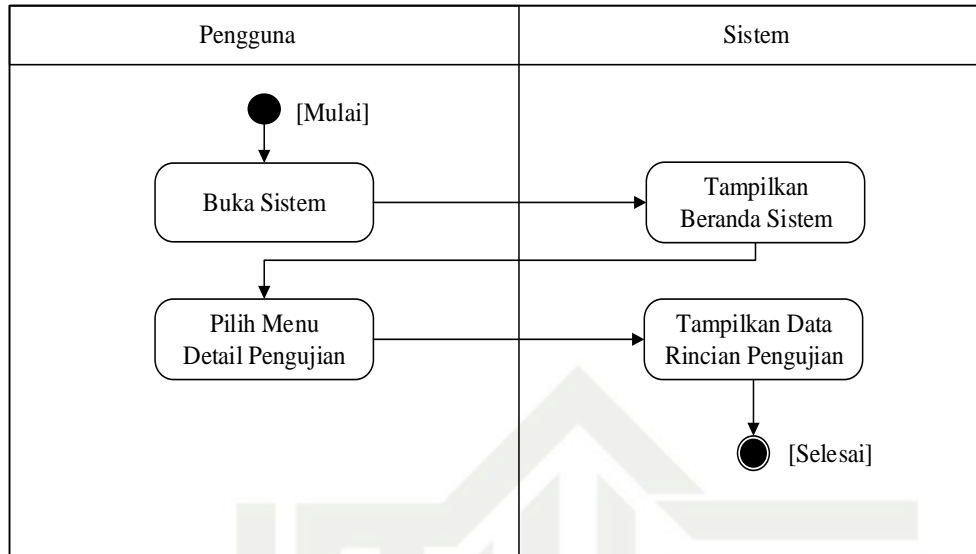
Gambar 4.6 Activity Diagram Pengujian

6. Lihat Detail Pengujian

Untuk *activity diagram* proses pada saat pengguna ingin melihat detail pengujian seperti pembagian data latih, data uji, nilai validitas dan *weight voting* pada sistem klasifikasi serangan jaringan ini dapat dilihat pada Gambar 4.7 berikut:

Hak Cipta Dilindungi Undang-Undang

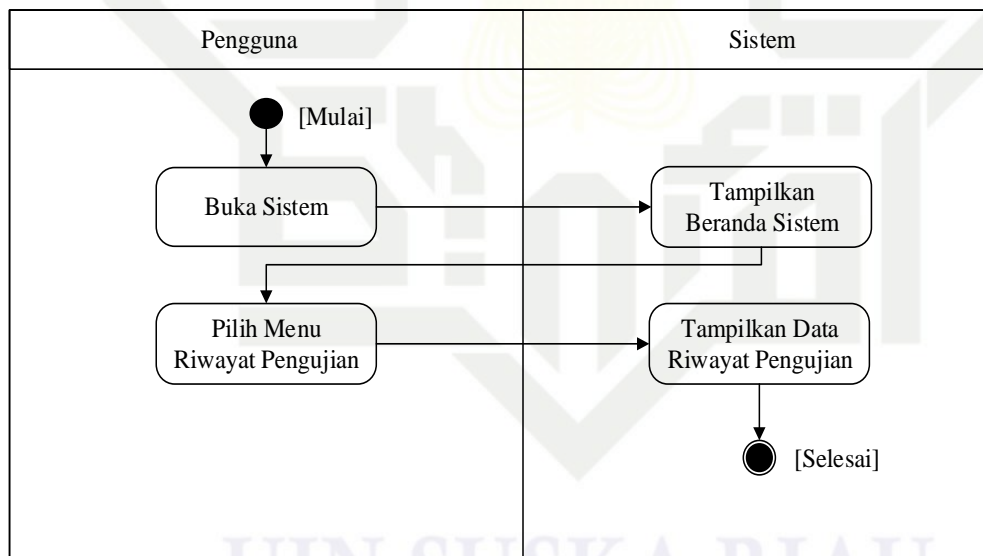
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan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



Gambar 4.7 Activity Diagram Lihat Detail Pengujian

7. Lihat Riwayat Pengujian

Untuk *activity diagram* proses pada saat pengguna ingin melihat riwayat beberapa pengujian yang telah dilakukan pada sistem klasifikasi serangan jaringan ini dapat dilihat pada Gambar 4.8 berikut:



Gambar 4.8 Activity Diagram Lihat Riwayat Pengujian

4.4.4 Sequence Diagram

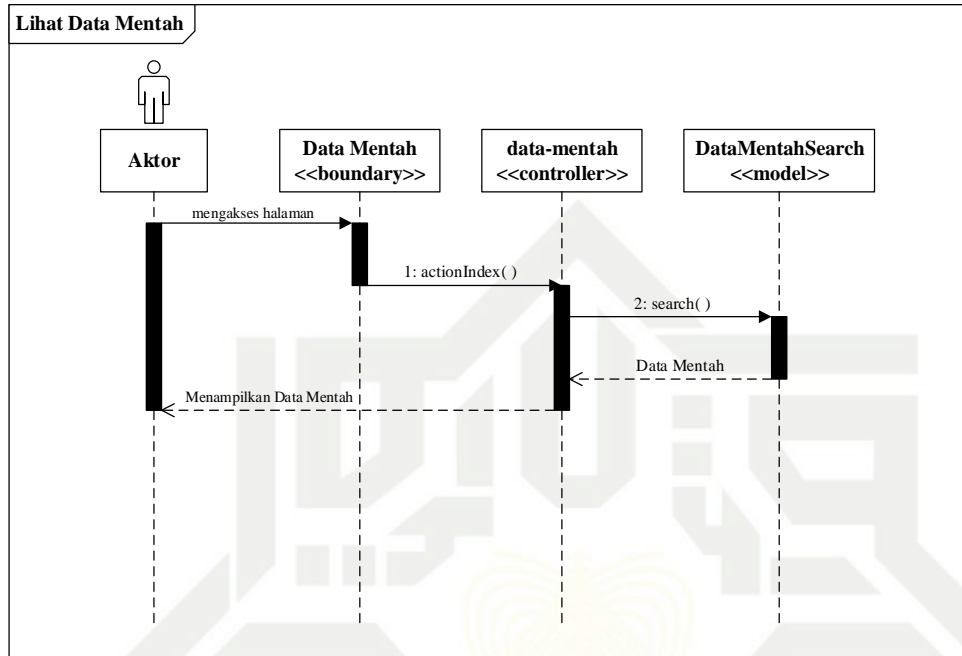
Sequence diagram akan memperlihatkan dan menampilkan beberapa interaksi antar objek di dalam sistem yang disusun pada sebuah urutan atau rangkaian waktu. *Sequence diagram* pada sistem klasifikasi serangan jaringan dapat dilihat pada Gambar 4.9 hingga Gambar 4.16 berikut:

Hak Cipta Diindungi Undang-Undang

1. Diarangi 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. Lihat Data Mentah

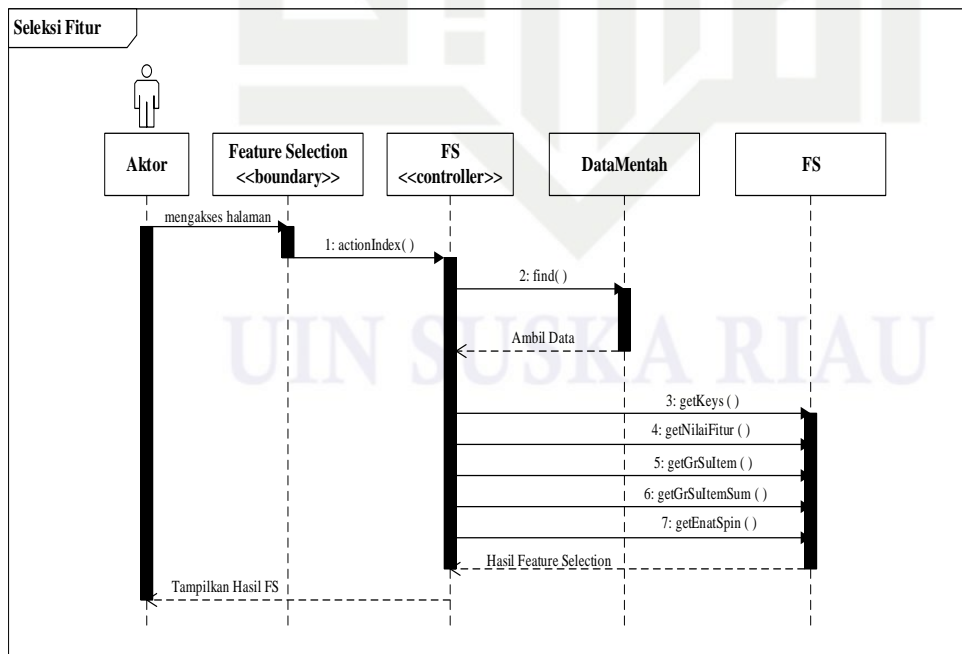
Untuk *sequence diagram* proses lihat data mentah pada sistem klasifikasi serangan jaringan ini dapat dilihat pada Gambar 4.9 berikut:



Gambar 4.9 *Sequence Diagram* Lihat Data Mentah

2. Seleksi Fitur

Untuk *sequence diagram* proses seleksi fitur pada sistem klasifikasi serangan jaringan ini dapat dilihat pada Gambar 4.10 berikut:



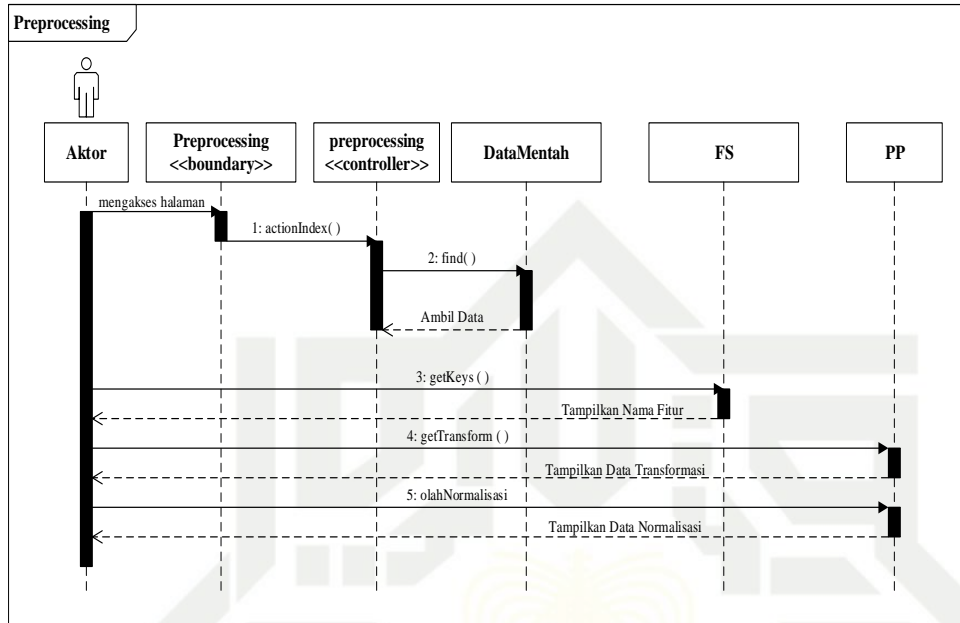
Gambar 4.10 *Sequence Diagram* Seleksi Fitur

Hak Cipta Diindungi 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. Preprocessing

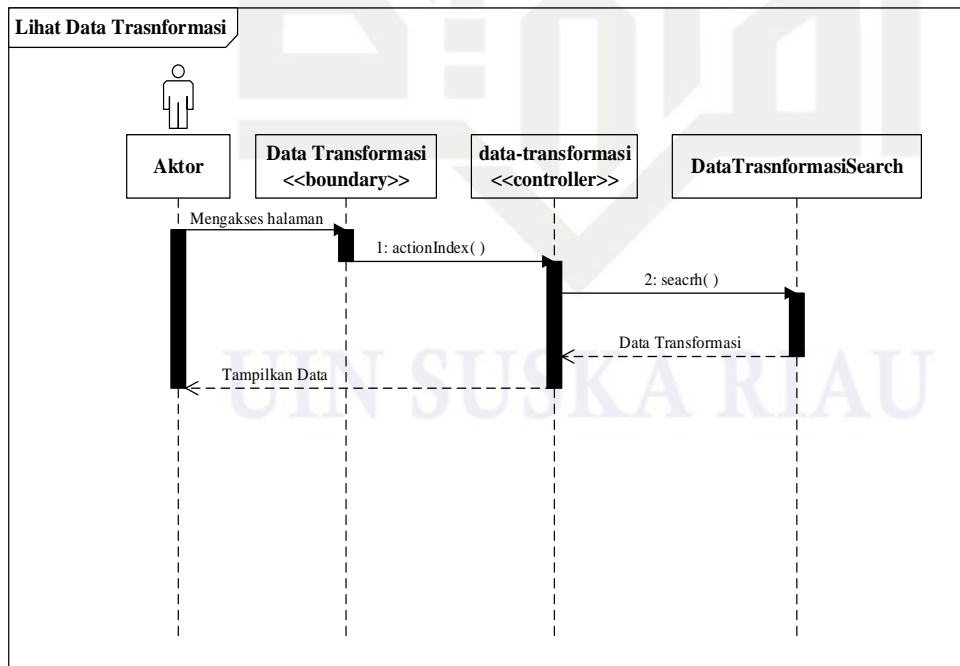
Untuk *sequence diagram* proses *preprocessing* pada sistem klasifikasi serangan jaringan ini dapat dilihat pada Gambar 4.11 berikut:



Gambar 4.11 *Sequence Diagram Preprocessing*

4. Lihat Data Transformasi

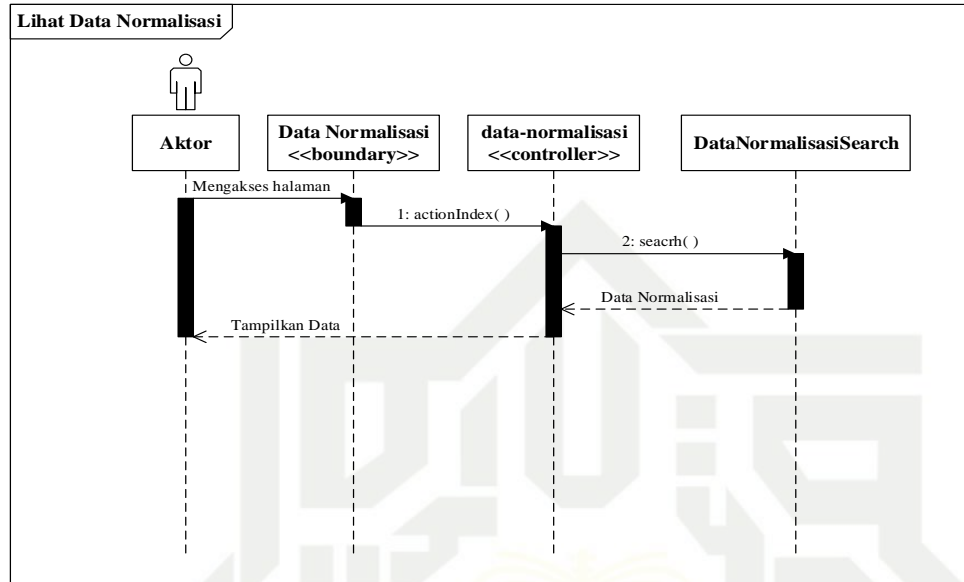
Untuk *sequence diagram* proses lihat data transformasi pada sistem klasifikasi serangan jaringan ini dapat dilihat pada Gambar 4.12 berikut:



Gambar 4.12 *Sequence Diagram Lihat Data Trasnformasi*

5. Lihat Data Normalisasi

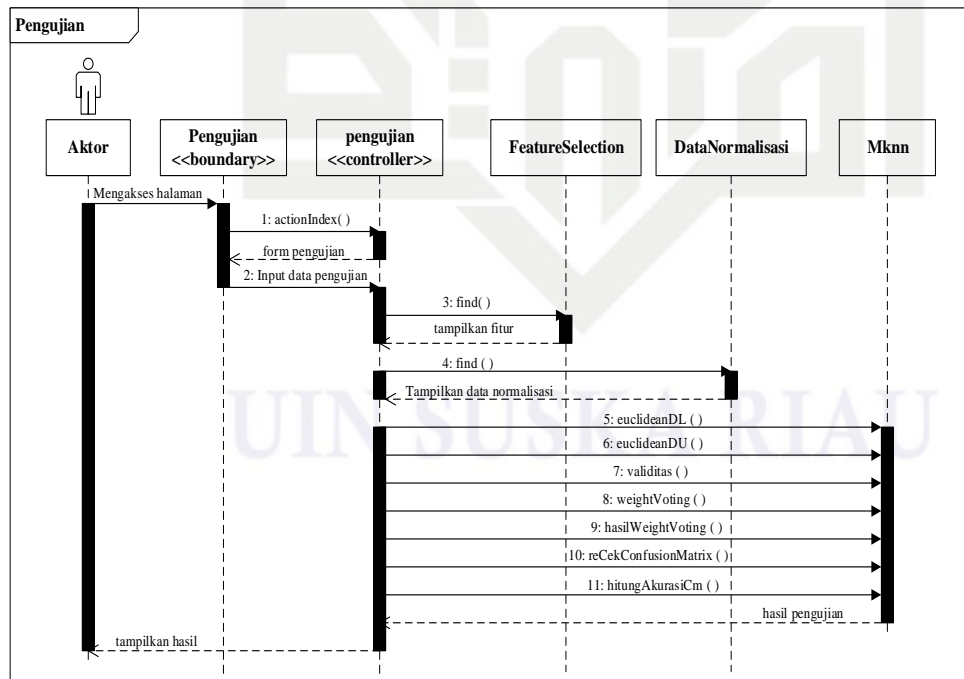
Untuk *sequence diagram* proses lihat data normalisasi pada sistem klasifikasi serangan jaringan ini dapat dilihat pada Gambar 4.13 berikut:



Gambar 4.13 Sequence Diagram Lihat Data Normalisasi

6. Pengujian

Untuk *sequence diagram* proses pengujian pada sistem klasifikasi serangan jaringan ini dapat dilihat pada Gambar 4.14 berikut:



Gambar 4.14 Sequence Diagram Pengujian

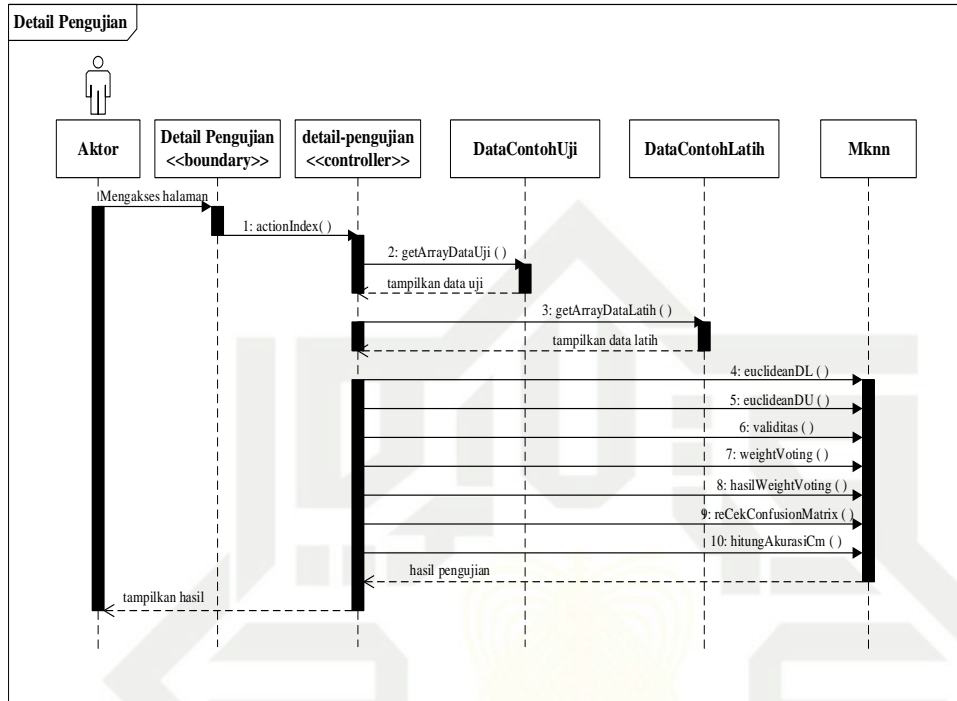
- Hak Cipta Diindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan 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 Diindungi 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. Lihat Detail Pengujian

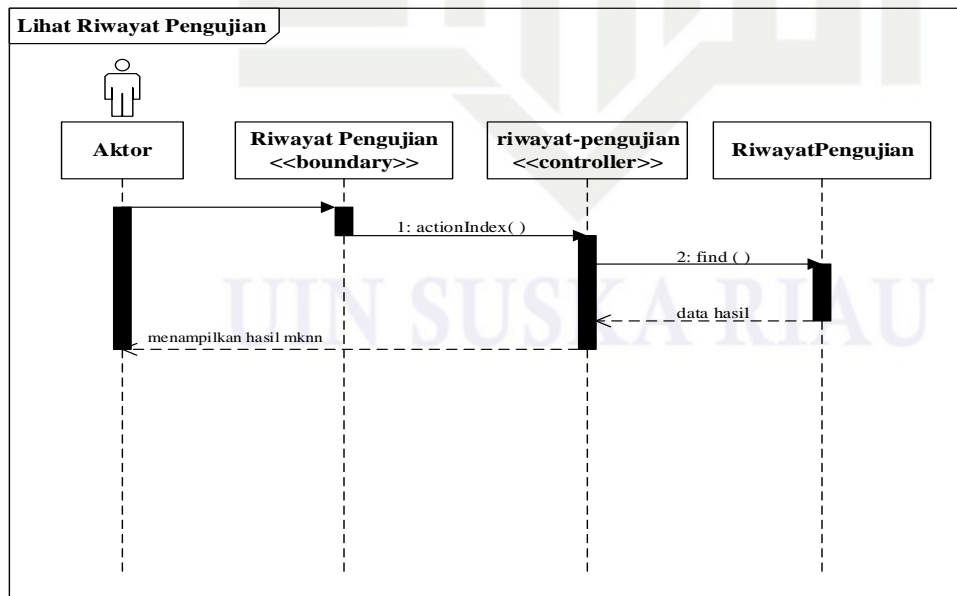
Untuk *sequence diagram* proses lihat detail pengujian pada sistem klasifikasi serangan jaringan ini dapat dilihat pada Gambar 4.15 berikut:



Gambar 4.15 *Sequence Diagram* Lihat Detail Pengujian

8. Lihat Riwayat Pengujian

Untuk *sequence diagram* proses lihat riwayat pengujian pada sistem klasifikasi serangan jaringan ini dapat dilihat pada Gambar 4.16 berikut:

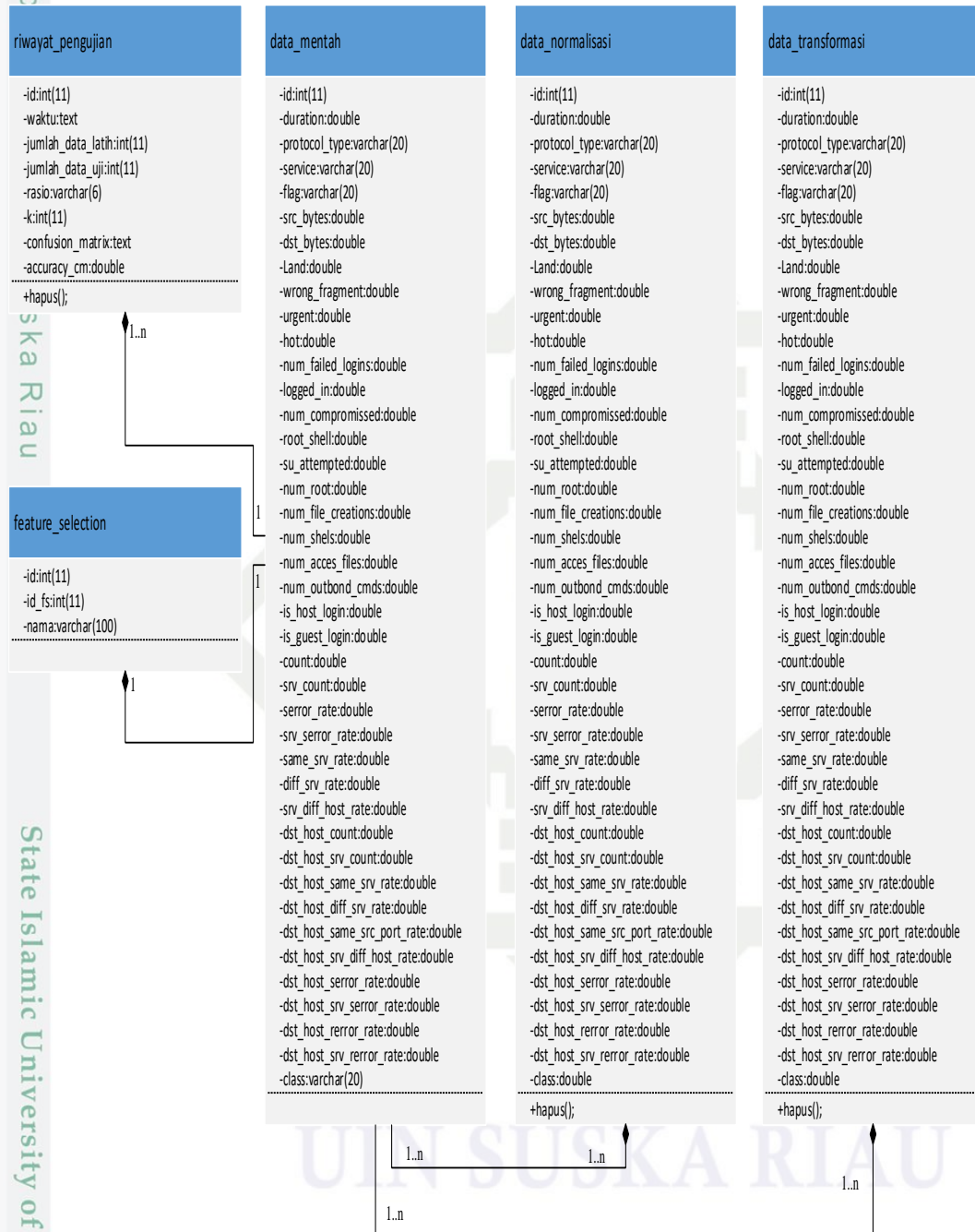


Gambar 4.16 *Sequence Diagram* Lihat Riwayat Pengujian

4.4.5 Class Diagram

Class diagram dari sistem yang dibangun dapat dilihat pada Gambar 4.17

berikut:

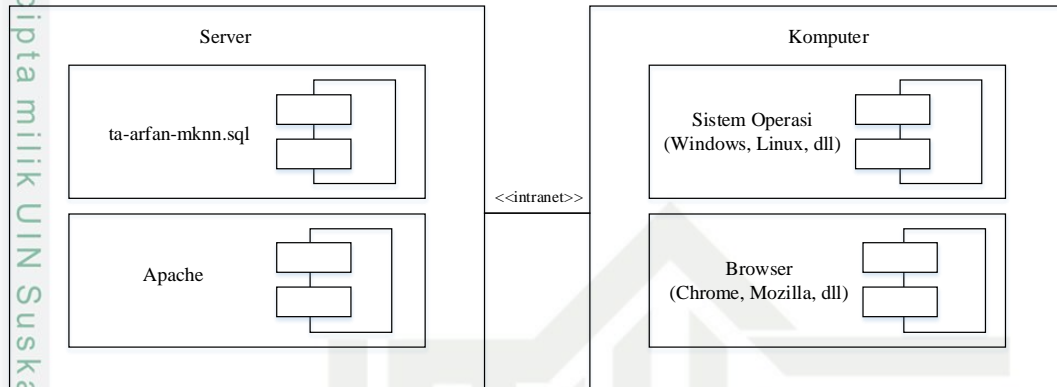


Gambar 4.17 Class Diagram

- Hak Cipta Diindungi 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.

4.4.6 Deployment Diagram

Deployment diagram dari sistem yang dibangun dapat dilihat pada Gambar 4.18 berikut:



Gambar 4.18 Deployment Diagram

4.5 Perancangan Database

Berikut merupakan perancangan table database pada penelitian ini.

4.4.1 Tabel Data Mentah

Nama Tabel : data_mentah

Deskripsi isi : Berisi data serangan jaringan pada dataset NSL-KDD

Primary key : *id*

Penjelasan struktur Tabel Data Mentah dapat dilihat pada Tabel 4.27.

Tabel 4.27 Struktur Tabel Data Mentah

No	Nama Field	Type Data	Deskripsi	Null	Key
1	Id	Int (11)	Nomor indeks serangan jaringan	Not Null	Primary Key (PK)
2	Duration	Double	Nilai dari serangan: (Duration)	Not Null	
3	Protocol_type	Varchar (20)	Nilai dari serangan: (protocol_type)	Not Null	
4	Service	Varchar (20)	Nilai dari serangan: (Service)	Not Null	
5	Flag	Varchar (20)	Nilai dari serangan: (flag)	Not Null	
6	Src_bytes	Double	Nilai dari serangan: (src_bytes)	Not Null	
7	Dst_bytes	Double	Nilai dari serangan: (dst_bytes)	Not Null	
8	Land	Double	Nilai dari serangan: (land)	Not Null	
9	Wrong_fragment	Double	Nilai dari serangan: (wrong_fragment)	Not Null	

No	Nama Field	Type Data	Deskripsi	Null	Key
10	Urgent	Double	Nilai dari serangan: (urgent)	Not Null	
11	Hot	Double	Nilai dari serangan: (hot)	Not Null	
12	Num_failed_login	Double	Nilai dari serangan: (num_failed_logins)	Not Null	
13	Logged_in	Double	Nilai dari serangan: (logged_in)	Not Null	
14	Num_compromised	Double	Nilai dari serangan: (num_compromised)	Not Null	
15	Root_shell	double	Nilai dari serangan: (root_shell)	Not Null	
16	Su_attempted	Double	Nilai dari serangan: (su_attempted)	Not Null	
17	Num_root	Double	Nilai dari serangan: (num_root)	Not Null	
18	Num_file_creations	Double	Nilai dari serangan: (num_file_creations)	Not Null	
19	Num_shells	Double	Nilai dari serangan: (num_shells)	Not Null	
20	Num_access_files	Double	Nilai dari serangan: (num_access_files)	Not Null	
21	Num_outbond_cmds	Double	Nilai dari serangan: (num_outbound_cmds)	Not Null	
22	Is_host_login	Double	Nilai dari serangan: (is_host_login)	Not Null	
23	Is_guest_login	Double	Nilai dari serangan: (is_guest_login)	Not Null	
24	Count	Double	Nilai dari serangan: (count)	Not Null	
25	Srv_count	Double	Nilai dari serangan: (srv_count)	Not Null	
26	Serror_rate	Double	Nilai dari serangan: (serror_rate)	Not Null	
27	Srv_serror_rate	Double	Nilai dari serangan: (srv_serror_rate)	Not Null	
28	Same_srv_rate	Double	Nilai dari serangan: (rerror_rate)	Not Null	
29	Diff_srv_rate	Double	Nilai dari serangan: (srv_rerror_rate)	Not Null	
30	Srv_diff_host_rate	Double	Nilai dari serangan: (same_srv_rate)	Not Null	
31	Dst_host_count	Double	Nilai dari serangan: (diff_srv_rate)	Not Null	
32	Dst_hot_srv_count	Double	Nilai dari serangan: (srv_diff_host_rate)	Not Null	
33	Dst_host_same_srv_rate	Double	Nilai dari serangan: (dst_host_count)	Not Null	
34	Dst_host_diff_rate	Double	Nilai dari serangan: (dst_host_srv_count)	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 Field	Tipe Data	Deskripsi	Null	Key
35	Dst_host_same_src_port_rate	Double	Nilai dari serangan: (dst_host_same_srv_rate)	Not Null	
36	Dst_host_diff_srv_rate	Double	Nilai dari serangan: (dst_host_diff_srv_rate)	Not Null	
37	Dst_host_same_src_port_rate	Double	Nilai dari serangan: (dst_host_same_src_port_rate)	Not Null	
38	Dst_host_srv_diff_host_rate	Double	Nilai dari serangan: (dst_host_srv_diff_host_rate)	Not Null	
39	Dst_host_error_rate	Double	Nilai dari serangan: (dst_host_error_rate)	Not Null	
40	Dst_host_srv_error_rate	Double	Nilai dari serangan: (dst_host_srv_error_rate)	Not Null	
41	Dst_host_error_rate	Double	Nilai dari serangan: (dst_host_error_rate)	Not Null	
42	Dst_host_srv_error_rate	Double	Nilai dari serangan: (dst_host_srv_error_rate)	Not Null	
43	Class	varchar (20)	Lima kelas jaringan yang terdiri dari Probe, R2L, U2L, Dos dan Normal.	Not Null	

4.5.7 Tabel Data Normalisasi

Nama Tabel : data_normalisasi

Deskripsi isi : Berisi data serangan dengan fitur yang telah terseleksi.

Primary key : **Id**

Penjelasan struktur Tabel Data Normalisasi dapat dilihat selengkapnya pada Tabel 4.28.

Tabel 4.28 Struktur Tabel Data Normalisasi

No	Nama Field	Tipe Dan Panjang Data	Deskripsi	Null	Key
1	Id	Int (11)	Berisi tentang data yang telah di normalisasi	Not Null	PK
2	Duration	Double	Berisi tentang data yang telah di normalisasi	Not Null	
3	Protocol_type	Varchar (20)	Berisi tentang data yang telah di normalisasi	Not Null	
4	Service	Varchar (20)	Berisi tentang data yang telah di normalisasi	Not Null	
5	Flag	Varchar (20)	Berisi tentang data yang telah di normalisasi	Not Null	

No	Nama Field	Tipe Dan Panjang Data	Deskripsi	Null	Key
6	Src_bytes	Double	Berisi tentang data yang telah di normalisasi	Not Null	
7	Dst_bytes	Double	Berisi tentang data yang telah di normalisasi	Not Null	
8	Land	Double	Berisi tentang data yang telah di normalisasi	Not Null	
9	Wrong_fragment	Double	Berisi tentang data yang telah di normalisasi	Not Null	
10	Urgent	Double	Berisi tentang data yang telah di normalisasi	Not Null	
11	Hot	Double	Berisi tentang data yang telah di normalisasi	Not Null	
12	Num_failed_login	Double	Berisi tentang data yang telah di normalisasi	Not Null	
13	Logged_in	Double	Berisi tentang data yang telah di normalisasi	Not Null	
14	Num_compromised	Double	Berisi tentang data yang telah di normalisasi	Not Null	
15	Root_shell	double	Berisi tentang data yang telah di normalisasi	Not Null	
16	Su_attempted	Double	Berisi tentang data yang telah di normalisasi	Not Null	
17	Num_root	Double	Berisi tentang data yang telah di normalisasi	Not Null	
18	Num_file_creations	Double	Berisi tentang data yang telah di normalisasi	Not Null	
19	Num_shells	Double	Berisi tentang data yang telah di normalisasi	Not Null	
20	Num_access_files	Double	Berisi tentang data yang telah di normalisasi	Not Null	
21	Num_outbond_cmds	Double	Berisi tentang data yang telah di normalisasi	Not Null	
22	Is_host_login	Double	Berisi tentang data yang telah di normalisasi	Not Null	
23	Is_guest_login	Double	Berisi tentang data yang telah di normalisasi	Not Null	
24	Count	Double	Berisi tentang data yang telah di normalisasi	Not Null	
25	Srv_count	Double	Berisi tentang data yang telah di normalisasi	Not Null	
26	Serror_rate	Double	Berisi tentang data yang telah di normalisasi	Not Null	
27	Srv_serror_rate	Double	Berisi tentang data yang telah di normalisasi	Not Null	
28	Same_srv_rate	Double	Berisi tentang data yang telah di normalisasi	Not Null	
29	Diff_srv_rate	Double	Berisi tentang data yang telah di normalisasi	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.

- 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 Field	Tipe Dan Panjang Data	Deskripsi	Null	Key
30	Srv_diff_host_rate	Double	Berisi tentang data yang telah di normalisasi	Not Null	
31	Dst_host_count	Double	Berisi tentang data yang telah di normalisasi	Not Null	
32	Dst_hot_srv_count	Double	Berisi tentang data yang telah di normalisasi	Not Null	
33	Dst_host_same_srv_rate	Double	Berisi Tentang Data Yang Telah Di Normalisasi	Not Null	
34	Dst_host_diff_rate	Double	Berisi tentang data yang telah di normalisasi	Not Null	
35	Dst_host_same_src_port_rate	Double	Berisi tentang data yang telah di normalisasi	Not Null	
36	Dst_host_diff_srv_rate	Double	Berisi tentang data yang telah di normalisasi	Not Null	
37	Dst_host_same_src_port_rate	Double	Berisi tentang data yang telah di normalisasi	Not Null	
38	Dst_host_srv_diff_host_rate	Double	Berisi tentang data yang telah di normalisasi	Not Null	
39	Dst_host_rerror_rate	Double	Berisi tentang data yang telah di normalisasi	Not Null	
40	Dst_host_srv_serror_rate	Double	Berisi tentang data yang telah di normalisasi	Not Null	
41	Dst_host_rerror_rate	Double	Berisi tentang data yang telah di normalisasi	Not Null	
42	Dst_host_srv_rerror_rate	Double	Berisi tentang data yang telah di normalisasi	Not Null	
43	Class	varchar (20)	Berisi tentang data yang telah di normalisasi	Not Null	

4.5.8 Tabel Data Transformasi

Nama Tabel : data_transformasi

Deskripsi isi : Berisi data serangan dengan fitur yang telah terseleksi.

Primary key : **Id**

Penjelasan struktur Tabel Data Transformasi dapat dilihat selengkapnya pada Tabel 4.29

Tabel 4.29 Struktur Tabel Data Transformasi

No	Nama Field	Tipe Dan Panjang Data	Deskripsi	Null	Key
1	<i>Id</i>	Int (11)	Berisi tentang data yang telah di transformasi	Not Null	PK

No	Nama Field	Tipe Dan Panjang Data	Deskripsi	Null	Key
2	<i>Duration</i>	<i>Double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
3	<i>Protocol_type</i>	<i>Varchar (20)</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
4	<i>Service</i>	<i>Varchar (20)</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
5	<i>Flag</i>	<i>Varchar (20)</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
6	<i>Src_bytes</i>	<i>Double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
7	<i>Dst_bytes</i>	<i>Double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
8	<i>Land</i>	<i>Double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
9	<i>Wrong_fragment</i>	<i>Double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
10	<i>Urgent</i>	<i>Double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
11	<i>Hot</i>	<i>Double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
12	<i>Num_failed_login</i>	<i>Double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
13	<i>Logged_in</i>	<i>Double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
14	<i>Num_compromised</i>	<i>Double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
15	<i>Root_shell</i>	<i>double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
16	<i>Su_attempted</i>	<i>Double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
17	<i>Num_root</i>	<i>Double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	

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 Field	Tipe Dan Panjang Data	Deskripsi	Null	Key
18	<i>Num_file_creations</i>	<i>Double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
19	<i>Num_shells</i>	<i>Double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
20	<i>Num_access_files</i>	<i>Double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
21	<i>Num_outbond_cmds</i>	<i>Double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
22	<i>Is_host_login</i>	<i>Double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
23	<i>Is_guest_login</i>	<i>Double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
24	<i>Count</i>	<i>Double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
25	<i>Srv_count</i>	<i>Double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
26	<i>Serror_rate</i>	<i>Double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
27	<i>Srv_serror_rate</i>	<i>Double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
28	<i>Same_srv_rate</i>	<i>Double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
29	<i>Diff_srv_rate</i>	<i>Double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
30	<i>Srv_diff_host_rate</i>	<i>Double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
31	<i>Dst_host_count</i>	<i>Double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
32	<i>Dst_hot_srv_count</i>	<i>Double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	
33	<i>Dst_host_same_srv_rate</i>	<i>Double</i>	Berisi tentang data yang telah di transformasi	<i>Not Null</i>	

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. 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 Field	Tipe Dan Panjang Data	Deskripsi	Null	Key
34	<i>Dst_host_diff_rate</i>	Double	Berisi tentang data yang telah di transformasi	Not Null	
35	<i>Dst_host_same_src_port_rate</i>	Double	Berisi tentang data yang telah di transformasi	Not Null	
36	<i>Dst_host_diff_srv_rate</i>	Double	Berisi tentang data yang telah di transformasi	Not Null	
37	<i>Dst_host_same_src_port_rate</i>	Double	Berisi tentang data yang telah di transformasi	Not Null	
38	<i>Dst_host_srv_diff_host_rate</i>	Double	Berisi tentang data yang telah di transformasi	Not Null	
39	<i>Dst_host_rerror_rate</i>	Double	Berisi tentang data yang telah di transformasi	Not Null	
40	<i>Dst_host_srv_serror_rate</i>	Double	Berisi tentang data yang telah di transformasi	Not Null	
41	<i>Dst_host_rerror_rate</i>	Double	Berisi tentang data yang telah di transformasi	Not Null	
42	<i>Dst_host_srv_rerror_rate</i>	Double	Berisi tentang data yang telah di transformasi	Not Null	
43	<i>Class</i>	varchar (20)	Berisi tentang data yang telah di transformasi	Not Null	

4.5.9 Tabel Feature Selection

Nama Tabel : *feature_selection*

Deskripsi isi : Berisi data serangan dengan fitur yang telah terseleksi.

Primary key : *Id*

Penjelasan struktur Tabel *Feature Selection* dapat dilihat selengkapnya pada Tabel

4.30

Tabel 4.30 Struktur Tabel Data *Feature Selection*

No	Nama Field	Tipe Dan Panjang Data	Deskripsi	Null	Key
1	<i>Id</i>	Int (11)	Nomor Indeks Seleksi Fitur	Not Null	PK

No	Nama Field	Tipe Dan Panjang Data	Deskripsi	Null	Key
2	<i>Id_fs</i>	<i>Int</i> (11)	Nilai Dari Fitur Terbaik 1	<i>Not Null</i>	
3	nama	<i>Varchar</i> (100)	Nilai Dari Fitur Terbaik 2	<i>Not Null</i>	

4.5.10 Tabel Riwayat Pengujian

Nama Tabel : riwayat_pengujian

Deskripsi isi : Berisi data serangan dengan fitur yang telah terseleksi.

Primary key : *Id*

Penjelasan struktur Tabel Data Riwayat Pengujian dapat dilihat selengkapnya pada Tabel 4.31.

Tabel 4.31 Struktur Tabel Data Riwayat Pengujian

No	Nama Field	Tipe Dan Panjang Data	Deskripsi	Null	Key
1	Id	<i>Int</i> (11)	Berisi tentang data pengujian	<i>Not Null</i>	PK
2	waktu	<i>Int</i> (11)	Berisi tentang data pengujian	<i>Not Null</i>	
3	Jumlah_data_latih	<i>Varchar</i> (100)	Berisi tentang data pengujian	<i>Not Null</i>	
4	Jumlah_data_uji	<i>Int</i> (11)	Berisi tentang data pengujian	<i>Not Null</i>	
5	rasio	<i>Int</i> (11)	Berisi tentang data pengujian	<i>Not Null</i>	
6	K	<i>Int</i> (11)	Berisi tentang data pengujian	<i>Not Null</i>	
7	<i>Confusion_matrix</i>	<i>Text</i>	Berisi tentang data pengujian	<i>Not Null</i>	
8	<i>Accuracy_cm</i>	<i>double</i>	Berisi tentang data pengujian	<i>Not Null</i>	

Hak Cipta Dilindungi Undang-Undang

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

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- 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 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.

4.5.11 Tabel Tema

Nama Tabel : tema
 Deskripsi isi : Berisi data serangan dengan fitur yang telah terseleksi.
 Primary key : *Id*

Penjelasan struktur Tabel Tema dapat dilihat selengkapnya pada Tabel 4.32

Tabel 4.32 Struktur Tabel Data Tema

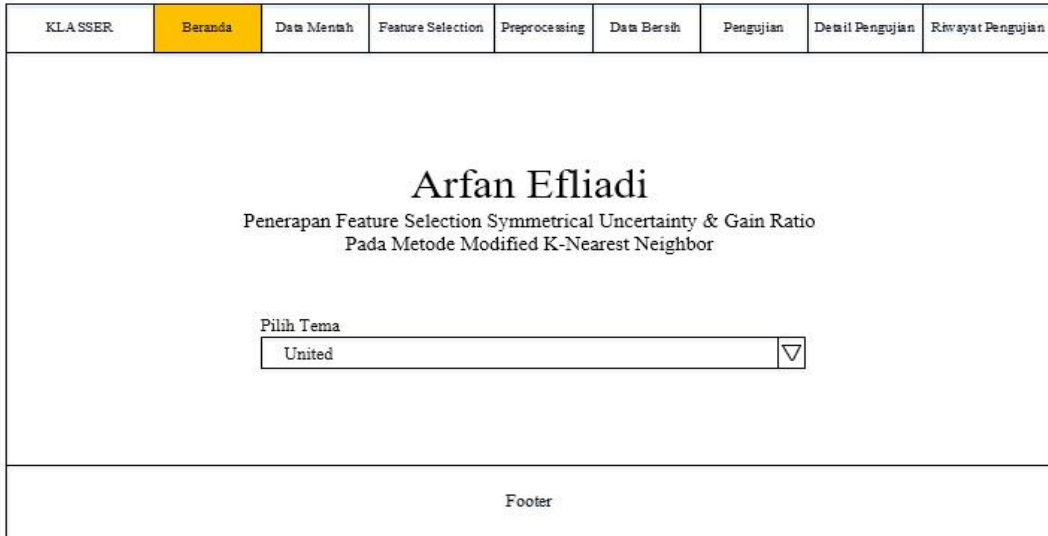
No	Nama Field	Tipe Dan Panjang Data	Deskripsi	Null	Key
1	<i>Id</i>	<i>Int (11)</i>	Berisi tentang data pengujian	<i>Not Null</i>	PK
2	nama	<i>Varchar (50)</i>	Berisi tentang data pengujian	<i>Not Null</i>	

4.6 Perancangan Antarmuka (*Interface*)

Interface system adalah sarana pengembangan sistem yang digunakan untuk membuat komunikasi lebih mudah, konsisten antara sistem dengan pemakainya. Penekanan *interface* tampilan yang baik dan mudah dipahami. Adapun struktur menu pada sistem ini dapat dilihat pada Gambar berikut.

4.5.1 Perancangan Halaman Beranda

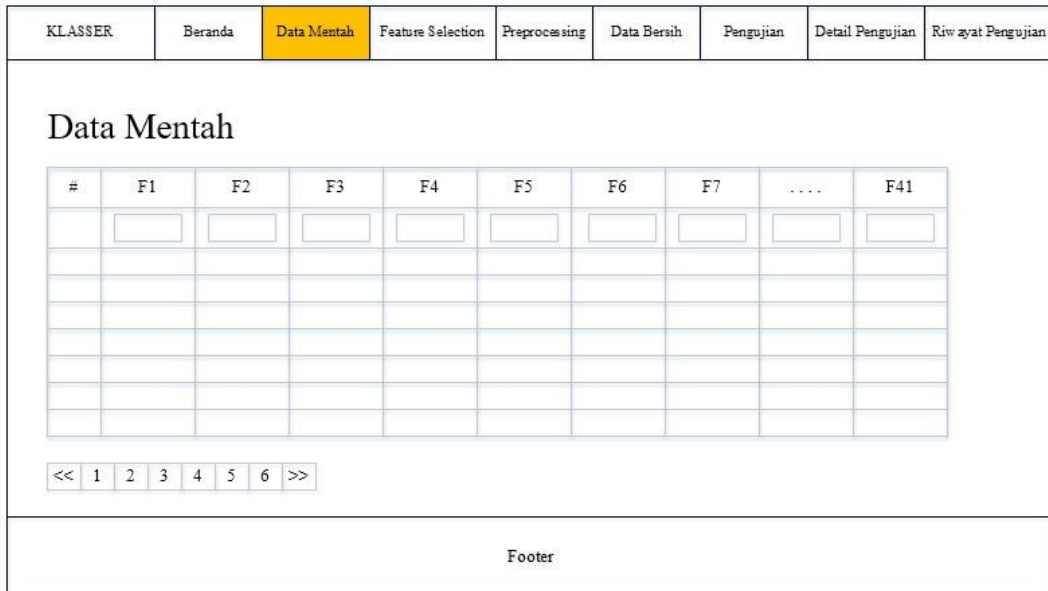
Halaman beranda adalah halaman paling awal ketika pengguna mengakses sistem. Pada halaman beranda ini terdapat sebuah menu untuk memilih tema, terdapat beberapa macam tema yang dapat digunakan. Perancangan halaman beranda dapat dilihat pada Gambar 4.19.



Gambar 4.19 Perancangan Halaman Beranda

4.6.12 Perancangan Halaman Data Mentah

Halaman Data Mentah adalah halaman yang berisi keseluruhan data yang digunakan pada penelitian ini, ketika pengguna membuka halaman ini secara langsung akan melihat data kotor sebelum dilakukan beberapa proses pembersihan data lainnya. Perancangan halaman data mentah dapat dilihat pada Gambar 4.20



Gambar 4.20 Perancangan Halaman Data Mentah

4.6.13 Perancangan Halaman *Feature Selection*

Halaman *feature selection* adalah halaman yang dapat diakses *user* untuk melakukan proses seleksi fitur terhadap data yang ada pada dataset NSL-KDD

- Hak Cipta Diindungi 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.

dengan menggunakan *Symmetrical Uncertainty* dan *Gain Ratio*. Perancangan halaman *feature selection* dapat dilihat pada Gambar 4.21

KLASSER	Beranda	Data Mentah	Feature Selection	Preprocessing	Data Bersih	Pengujian	Detail Pengujian	Riwayat Pengujian																																																																																																			
<div style="display: flex; justify-content: space-between;"> <div style="width: 20%;"> <p>Feature Selection</p> <ul style="list-style-type: none"> SUM FS SU & GR Perangkingan Boolean AND HASIL FS </div> <div style="width: 75%;"> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th colspan="9">duration</th> </tr> <tr> <th>Partisi</th> <th>Jumlah Kasus</th> <th>Normal</th> <th>Dos</th> <th>Probe</th> <th>R2L</th> <th>Entropy</th> <th>Entropy Atribut</th> <th>Split Info</th> </tr> </thead> <tbody> <tr><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td></tr> <tr><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td></tr> <tr><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td></tr> <tr><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td></tr> <tr><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td></tr> <tr><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td></tr> <tr><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td></tr> <tr><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td></tr> <tr><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td></tr> </tbody> </table> </div> </div> <div style="text-align: center; margin-top: 10px;"> << 1 2 3 4 5 6 >> </div>									duration									Partisi	Jumlah Kasus	Normal	Dos	Probe	R2L	Entropy	Entropy Atribut	Split Info																																																																																	
duration																																																																																																											
Partisi	Jumlah Kasus	Normal	Dos	Probe	R2L	Entropy	Entropy Atribut	Split Info																																																																																																			
Footer																																																																																																											

Gambar 4.21 Perancangan Halaman *Feature Selection*

4.6.14 Perancangan Halaman *Preprocessing*

Halaman *preprocessing* adalah halaman untuk memproses data mentah agar dibersihkan melalui proses transformasi data dan normalisasi, ketika *user* membuka halaman ini akan menerima informasi berupa data telah berhasil di transformasikan dan dinormalisasikan. Perancangan halaman *preprocessing* dapat dilihat pada Gambar 4.22

KLASSER	Beranda	Data Mentah	Feature Selection	Preprocessing	Data Bersih	Pengujian	Detail Pengujian	Riwayat Pengujian
<div style="background-color: #d4edda; padding: 10px; border: 1px solid #c3e6cb; margin-bottom: 10px;"> <p>Data berhasil di transformasi & dinormalisasi X</p> </div> <div style="border: 1px solid #ccc; padding: 10px; margin-bottom: 10px;"> <p>Buka menu "Data Bersih" untuk melihat bentuk data setelah preprocessing</p> </div>								
Footer								

Gambar 4.22 Perancangan Halaman *Preprocessing*

Hak Cipta Diindungi 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 Dilindungi Undang-Undang

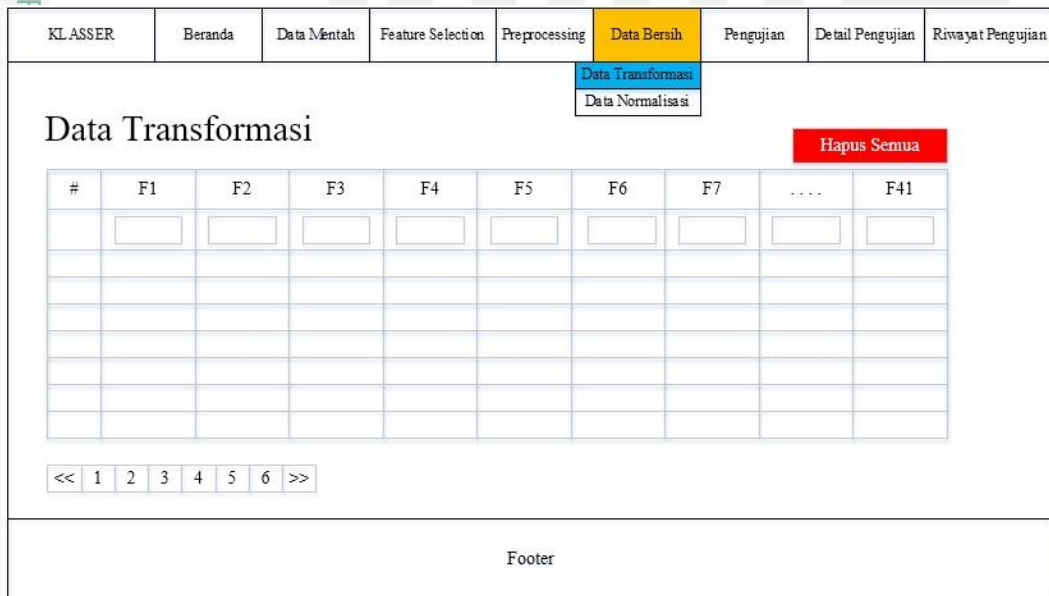
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan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

4.6.15 Perancangan Halaman Data Bersih

Halaman *data bersih* adalah halaman yang berisi data yang telah melalui proses transformasi data dan normalisasi, terdapat menu *dropdown* pada halaman ini yaitu halaman data transformasi dan halaman data normalisasi. Perancangan halaman data transformasi dan data normalisasi dapat dilihat pada Gambar 4.23 dan 4.24 berikut:

4.6.16 Perancangan Halaman Data Transformasi

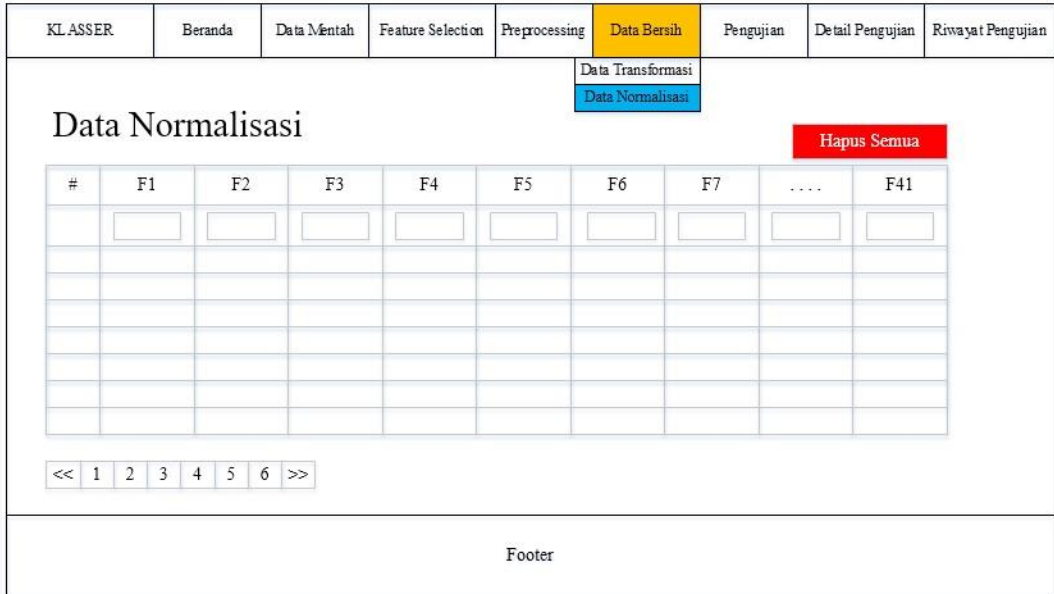
Halaman *data transformasi* berisi beberapa data yang telah dilakukan tahap transformasi. Perancangan halaman data transformasi dapat dilihat pada Gambar 4.23



Gambar 4.23 Perancangan Halaman Data Transformasi

4.6.17 Perancangan Halaman Data Normalisasi

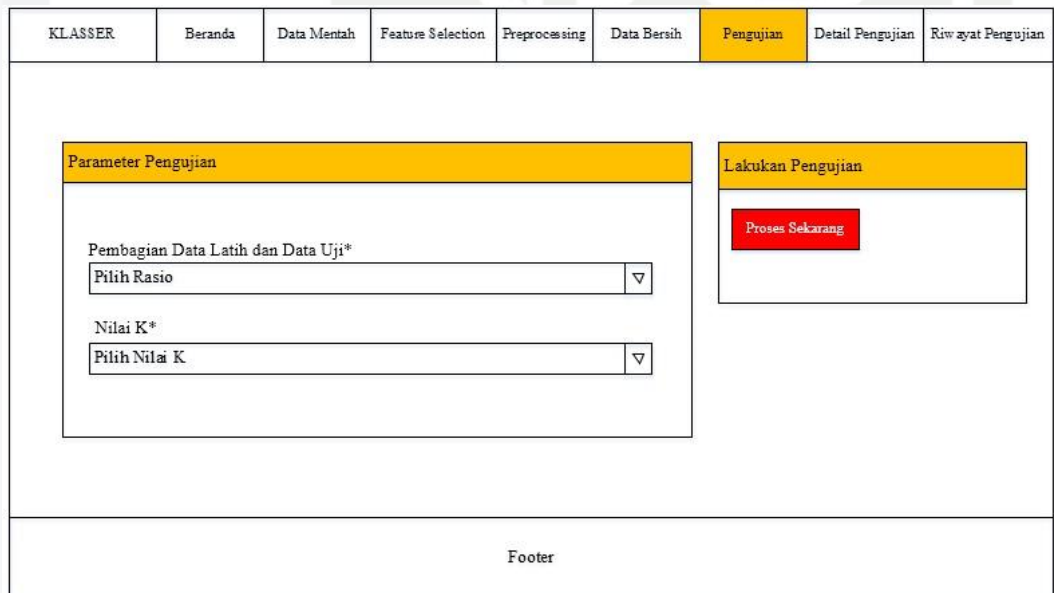
Halaman data normalisasi berisi tentang beberapa data yang telah melalui tahap normalisasi data. Perancangan halaman data normalisasi dapat dilihat pada Gambar 4.24



Gambar 4.24 Perancangan Halaman Data Normalisasi

4.6.18 Perancangan Halaman Pengujian

Halaman pengujian adalah halaman yang bisa digunakan *user* untuk melakukan pengujian terhadap nilai akurasi dan kinerja sistem. Terdapat beberapa tombol pada halaman ini seperti tombol untuk memilih rasio pembagian data latih dan uji, nilai K, dan tombol untuk memproses pengujian. Perancangan halaman pengujian dapat dilihat pada Gambar 4.25



Gambar 4.25 Perancangan Halaman Pengujian

- 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.

4.6.19 Perancangan Halaman Riwayat Pengujian

Halaman riwayat pengujian adalah halaman yang berisi beberapa riwayat pengujian yang telah dilakukan serta terdapat beberapa tombol *shortcut* untuk melihat kembali *confusion matrix* dan juga dapat menghapus riwayat pengujian. Perancangan halaman beranda dapat dilihat pada Gambar 4.26

KLASSER	Beranda	Data Mentah	Feature Selection	Preprocessing	Data Bersih	Pengujian	Detail Pengujian	Riwayat Pengujian
<div style="float: right;"> Hapus Semua Search: <input type="text"/> </div>								
No	Waktu Pengujian	Rasio	Jumlah Data Latih	Jumlah Data Uji	Nilai K	Confusion Matrix	Accuracy	Aksi
						Lihat Confusion Matrix		Hapus
Footer								

Gambar 4.26 Perancangan Halaman Riwayat Pengujian