

BAB III

ANALISIS dan PERANCANGAN SISTEM

3.1 Analisis Sistem

Berdasarkan hasil dari wawancara yang dilakukan di SMA Semen Gresik dalam menentukan jurusan siswa memerlukan beberapa data diantaranya adalah data hasil psikotes data minat jurusan dan data nilai *raport* siswa kelas IX. Nilai *raport* siswa yang digunakan untuk menentukan jurusan siswa adalah nilai *raport* Bahasa Indonesia, *raport* Bahasa Inggris, nilai *raport* matematika, nilai *raport* IPA, dan nilai *raport* IPS. Setelah memperoleh data nilai *raport* siswa psikotes serta minat siswa maka guru BK dapat menentukan jurusan siswa dengan melakukan pengelompokan dan menghitung seluruh data yang telah diperoleh hasil dari penentuan jurusan yang dilakukan setelah itu diserahkan kepada kepala sekolah untuk mendapat persetujuan.

Penentuan jurusan di SMA Semen Gresik masih dilakukan dengan mengelompokan data nilai *raport* siswa, minat dan hasil psikotes serta banyaknya data siswa setiap tahun membuat guru BK merasa kesulitan dalam menentukan jurusan siswa, ditambah lagi semakin banyak siswa maka semakin lama pula waktu yang diperlukan dalam menentukan jurusan siswa yang akan berdampak dalam kegiatan belajar.

3.2 Hasil Analisis

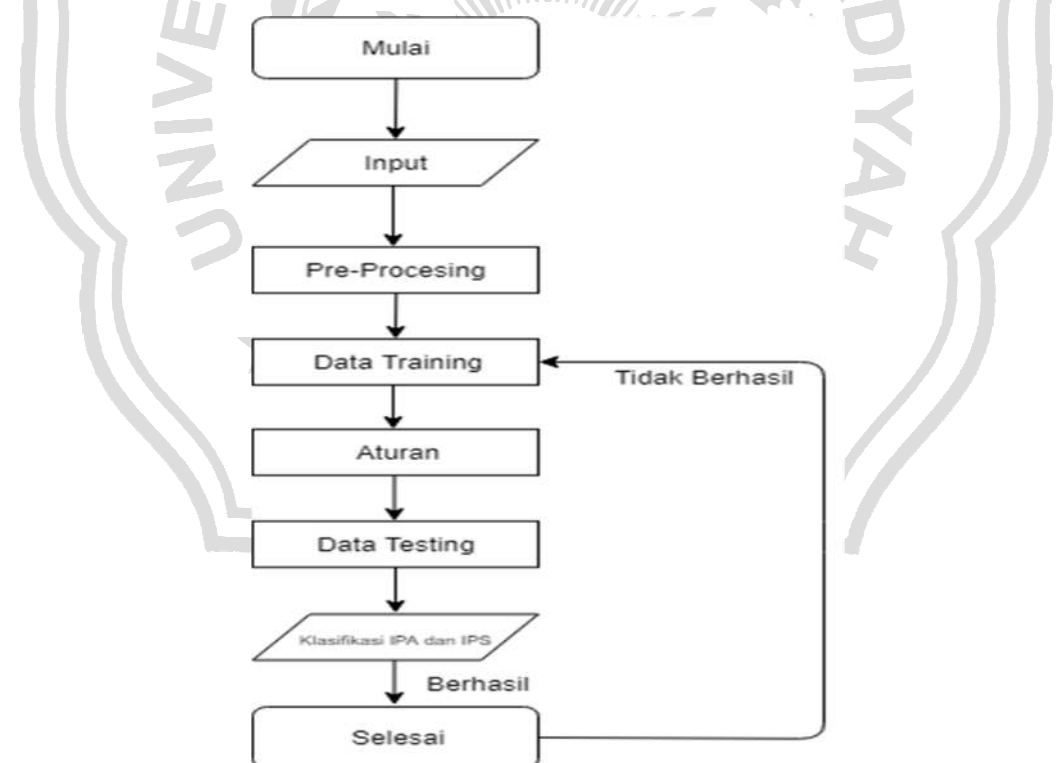
Sistem penentuan jurusan siswa merupakan suatu aplikasi data *mining* yang dapat membantu dalam menentukan jurusan siswa di SMA Semen Gresik. Penentuan jurusan siswa ini berguna untuk mengetahui jurusan yang tepat untuk setiap siswa dengan nilai *raport*, minat, serta psikotes sehingga dapat membantu guru BK dalam menentukan jurusan sesuai dengan kriteria yang ditentukan.

Nilai *raport* data minat dan data hasil psikotes tahun ajaran 2018-2019 akan digunakan sebagai data acuan dalam menghitung nilai rata-rata minat dan hasil psikotes sebagai nilai kriteria kemudian data tersebut akan direkap terlebih dahulu sebelum dilakukan proses perhitungan klasifikasi menggunakan metode *Decision Tree C45*.

Sistem yang akan dibangun ditujukan untuk guru BK sehingga dapat membantu dalam menentukan jurusan IPA atau IPS yang sesuai dengan kriteria yang sudah ditentukan. Dalam sistem ini terdapat 2 (dua) entitas, yaitu:

- a. Guru BK yang bertugas memasukkan data nilai kriteria.
- b. Kepala sekolah merupakan entitas yang hanya dapat melihat laporan hasil penjurusan siswa.

Kerja sistem diawali dengan menghitung nilai probabilitas masing-masing fitur dan kelas dari data latih. Selanjutnya sistem akan menghitung nilai probabilitas akhir (*posterior probability*) data uji terhadap data latih pada masing-masing kelas. Nilai probabilitas terbesar akan menentukan kategori kelas dari data yang diujikan. Hasil dari sistem ini nantinya adalah menampilkan rekomendasi jurusan yang sesuai dengan kriteria untuk siswa SMA Semen Gresik. Berikut ini adalah *FlowChart* Penentuan Penjurusan siswa SMA Semen Gresik pada gambar 3.1.



Gambar 3. 1 Flowchart System Penjurusan SMA

Penjelasan gambar 3.1:

1. Proses dimulai
 2. *Input* data siswa yang akan di proses
 3. Pre-processing data dilakukan dengan mentransformasikan data atribut yang bersifat numerik ke kategorikal serta membagi data yang akan diolah menjadi 2 kelompok yaitu data *training* dan data *testing*
 4. Data *training* dilakukan sbagai pemodelan pembelajaran data *mining* untuk mendapatkan pohon keputusan dan *rule* yang trbntuk
 5. Atursan berhasil terbntuk dari proses permodelan data *training* yang smpurna
 6. Data *testing* dilakukan yaitu dengan menguji proses penjurusan siswa dengan aturan yang terbentuk dari data *testing*
 7. Klasifikasi IPA dan IPS merupakan hasil yang diharapkan Jika proses tidak berhasil maka mengulangi lagi proses data *training* jika berhasil maka proses selesai
- Proses selesai

3.3 Kebutuhan Data

Tahapan awal yang dilakukan dalam penelitian ini adalah menyiapkan data data yang digunakan adalah periode tahun 2018-2019 data yang diperoleh dalam penelitian ini berupa nilai *raport* kelas IX untuk mata pelajaran (B Indonesia, B Inggris, Matematika, IPA, dan IPS) nilai minat, serta hasil psikotes. Jumlah data yang digunakan sbanyak 71 record dengan kelas IPA dan IPS masing masing berjumlah 30 dan 41 yang akan dibagi menjadi menjadi data latih dan data uji.

3.4 Persiapan Data

Dari 71 data diambil 30 data yang akan dijadikan data latih dan 41 sebagai data uji. Data latih disajikan pada tabel 3.1. Sedangkan data uji disajikan pada tabel 3.2.

Tabel 3. 1 Data Latih

NO	BINDO	BINGG	MTK	IPA	IPS	IQ	MINAT	HASIL
1	82	83	78	83	78	94	IPS	IPA
2	88	84	85	84	85	106	IPS	IPA
3	87	90	86	94	86	113	IPA	IPA
4	87	91	87	94	87	94	IPA	IPA
5	86	84	81	81	81	86	IPS	IPA
6	78	79	79	80	79	81	IPA	IPA
7	89	85	84	82	84	91	IPA	IPA
8	84	80	78	80	78	98	IPS	IPA
9	86	90	85	82	85	84	IPS	IPA
10	78	86	76	77	76	81	IPS	IPA
11	89	84	88	83	88	91	IPS	IPA
12	80	82	81	79	81	113	IPS	IPA
13	91	83	83	82	83	91	IPS	IPA
14	87	82	78	81	78	94	IPS	IPA
15	89	85	85	82	85	91	IPS	IPA
16	87	83	82	82	82	76	IPS	IPS
17	85	83	82	83	82	73	IPS	IPS
18	80	88	76	78	76	94	IPS	IPS
19	77	81	76	77	76	94	IPS	IPS
20	77	78	78	75	78	81	IPS	IPS
21	87	85	82	83	82	89	IPS	IPS
22	87	87	83	87	83	106	IPA	IPS
23	85	79	81	83	81	91	IPS	IPS
24	79	83	77	78	77	84	IPA	IPS
25	83	81	81	80	81	76	IPS	IPS
26	88	94	86	84	86	113	IPS	IPS
27	84	84	80	80	80	91	IPS	IPS
28	83	81	83	85	83	91	IPA	IPS
29	87	83	82	84	82	81	IPS	IPS
30	77	80	80	80	80	66	IPA	IPS

Tabel 3. 2 Data Uji

NO	BINDO	BINGG	MTK	IPA	IPS	IQ	MINAT	HASIL
1	80	79	82	79	82	87	IPS	IPA
2	92	85	84	82	84	91	IPS	IPA
3	86	88	90	85	90	98	IPS	IPA
4	90	88	88	85	88	106	IPS	IPA
5	81	81	78	82	78	98	IPS	IPA
6	85	81	77	79	77	84	IPS	IPA
7	84	79	78	79	78	98	IPA	IPA
8	83	83	83	80	83	76	IPS	IPA
9	88	83	82	81	82	84	IPS	IPA
10	83	83	87	84	87	106	IPS	IPA
11	78	79	78	78	78	76	IPS	IPA
12	85	85	90	79	90	94	IPS	IPA
13	86	87	82	85	82	98	IPA	IPA
14	84	84	81	81	81	81	IPS	IPA
15	84	83	81	80	81	73	IPS	IPA
16	82	84	82	81	82	76	IPS	IPA
17	86	87	88	88	88	101	IPA	IPA
18	84	81	83	81	83	84	IPS	IPA
19	80	79	77	77	79	88	IPA	IPA
20	91	87	84	87	84	94	IPS	IPA
21	88	91	85	89	85	94	IPA	IPA
22	89	84	84	81	84	87	IPS	IPA
23	80	90	78	87	78	109	IPS	IPA
24	81	81	80	81	80	73	IPA	IPA
25	88	84	81	83	81	106	IPA	IPA
26	81	78	80	83	80	73	IPA	IPA
27	90	90	84	88	84	98	IPS	IPA
28	83	86	81	81	81	73	IPS	IPA
29	81	80	83	81	84	98	IPS	IPA
30	89	86	82	81	82	84	IPA	IPA
31	79	79	79	78	79	91	IPS	IPA
32	84	85	81	82	81	98	IPS	IPA
33	90	82	84	82	84	84	IPS	IPA
34	88	85	85	84	85	84	IPS	IPA
35	84	84	82	79	82	87	IPS	IPA
36	82	82	82	80	82	94	IPS	IPA
37	84	80	83	85	83	73	IPS	IPS
38	83	81	79	80	79	84	IPA	IPS

39	87	83	80	83	80	64	IPA	IPS
40	83	82	81	82	81	98	IPS	IPS
41	79	80	80	81	80	71	IPA	IPS

Tabel 3. 3 Data Atribut

NO	ATRIBUT	KETERANGAN	TIPE
1	B INDO	NILAI <i>RAPORT</i>	NUMERIK
2	B INGGRIS	NILAI <i>RAPORT</i>	NUMERIK
3	MTK	NILAI <i>RAPORT</i>	NUMERIK
4	IPA	NILAI <i>RAPORT</i>	NUMERIK
5	IPS	NILAI <i>RAPORT</i>	NUMERIK
6	IQ	HASIL PSIKOTES	NUMERIK
7	MINAT	MINAT SISWA	NON NUMERIK

3.5 Perhitungan Decision Tree C45

Perhitungan *Decision Tree C4.5* ini akan menggunakan data pada tabel 3.1 (data *training*). Tabel tersebut akan diubah menjadi sebuah pohon keputusan. Sebelum melakukan perhitungan, berikut akan dijelaskan ketentuan dalam pembentukan pohon keputusan pada kasus ini:

- a) Pemecahan cabang dilakukan secara biner yaitu pemecahan yang hanya mempunyai dua nilai yakni \leq dan $>$ (kurang dari sama dengan dan lebih dari)
- b) Bahasa Indonesia (*Raport B Indo*)
Posisi v yang digunakan pada atribut *Raport B Indo* adalah nilai antara (77, 82 ,87, ,92)
- c) Bahasa Inggris (*Raport B Inggris*)
Posisi v yang digunakan pada atribut *Raport B Inggris* adalah nilai antara (77, 82 ,87, ,92)
- d) Matematika (*Raport MTK*)
Posisi v yang digunakan pada atribut *Raport MTK* adalah nilai antara (77, 82 ,87,92)

e) Ilmu Pengetahuan Alam (*Raport* IPA)

Posisi v yang digunakan pada atribut *Raport* IPA adalah nilai antara (77, 82, 87, 92).

f) Ilmu Pengetahuan Sosial (*Raport* IPS)

Posisi v yang digunakan pada atribut *Raport* IPS adalah nilai antara (77, 82, 87, 92).

g) IQ (Hasil Psikotes)

Posisi v yang digunakan pada atribut IQ adalah nilai antara (77, 82, 87, 92).

h) Minat

Pada variabel diagnosa yang Minat klasifikasi yang digunakan adalah IPA dan IPS.

Langkah pertama adalah memilih atribut yang akan dijadikan akar (*root node*) dengan menghitung nilai *gain* yang paling tinggi. Sebelumnya yang akan dihitung adalah nilai *entropy* semua data. Berikut adalah perhitungan *entropy* semua data.

$$\begin{aligned} Entropy(S) &= -\frac{30}{30} * \log_2\left(\frac{30}{30}\right) - \frac{30}{30} * \log_2\left(\frac{30}{30}\right) \dots\dots\dots(2.5) \\ &= 0,5 + 0,5 = 1 \end{aligned}$$

Selanjutnya menghitung nilai *gain* untuk setiap atribut, jika atribut memiliki nilai *numeric* maka akan dilakukan perhitungan untuk menentukan posisi v . Berikut adalah perhitungan untuk menentukan posisi *gain* tertinggi untuk masing-masing posisi. Hasilnya disajikan pada tabel 3.4

Tabel 3. 4 Posisi v untuk Pemecahan Atribut *Raport* B indo

<i>Raport</i> B Indo	77		82		87		92	
	<=	>	<=	>	<=	>	<=	>
IPA	0	15	4	11	10	5	15	0
IPS	3	12	5	10	17	1	15	0
JUMLAH	3	27	9	21	27	6	30	0
ENTROPY	0,0000	0,9911	0,9911	0,9984	0,9510	0,6500	1,0000	0,0000
GAIN	0,1080		0,0038		0,0141		0,0000	

Hasil penentuan posisi v pada atribut *Raport B Indo* menunjukkan nilai *gain* tertinggi didapat pada posisi $v=77$. Maka untuk atribut *Raport B Indo* dilakukan diskretisasi pada $v=77$ ketika menghitung *entropy* dan *gain* pada semua atribut. Hasilnya disajikan pada tabel 3.5.

Tabel 3. 5 Posisi v untuk Pemecahan Atribut *Raport B Ingg*

<i>Raport B Inggris</i>	77		82		87		92	
	\leq	$>$	\leq	$>$	\leq	$>$	\leq	$>$
IPA	0	15	4	11	12	3	15	0
IPS	0	15	6	9	13	2	14	1
JUMLAH	0	30	10	20	25	5	29	1
<i>ENTROPY</i>	0,0000	1,0000	0,9710	0,9928	0,9988	0,9710	0,9991	0,0000
<i>GAIN</i>	0,0000		0,0145		0,0058		0,0342	

Hasil penentuan posisi v pada atribut *Raport B Inggris* menunjukkan nilai *gain* tertinggi didapat pada posisi $v=92$. Maka atribut *Raport B Inggris* dilakukan diskretisasi pada $v=92$ ketika menghitung *entropy* dan *gain* pada semua atribut. Hasilnya disajikan pada tabel 3.6.

Tabel 3. 6 Posisi v untuk Pemecahan Atribut *Raport Mtk*

<i>Raport MTK</i>	77		82		87		92	
	\leq	$>$	\leq	$>$	\leq	$>$	\leq	$>$
IPA	1	14	7	8	14	1	15	0
IPS	3	12	12	3	15	0	15	0
JUMLAH	4	26	19	11	29	1	30	0
<i>ENTROPY</i>	0,8113	0,9957	0,9495	0,8454	0,9991	0,0000	1,0000	0,0000
<i>GAIN</i>	0,0289		0,0887		0,0342		0,0000	

Hasil penentuan posisi v pada atribut *Raport MTK* menunjukkan nilai *gain* tertinggi didapat pada posisi $v=82$. Maka atribut *Raport MTK* dilakukan diskretisasi pada $v=82$ ketika menghitung *entropy* dan *gain* pada semua atribut. Hasilnya disajikan pada tabel 3.7.

Tabel 3. 7 Posisi v untuk Pemecahan Atribut *Raport* IPA

<i>Raport</i> IPA	77		82		87		92	
	\leq	$>$	\leq	$>$	\leq	$>$	\leq	$>$
IPA	1	14	10	5	13	2	13	2
IPS	2	13	8	7	15	0	15	0
JUMLAH	3	27	18	12	28	2	28	2
ENTROPY	0,9183	0,9990	0,9911	0,9799	0,9963	0,0000	0,9963	0,0000
GAIN	0,0091		0,0134		0,0701		0,0701	

Hasil penentuan posisi v pada atribut *Raport* IPA menunjukkan nilai *gain* tertinggi didapat pada posisi $v=87$. Maka untuk atribut *Raport* IPA dilakukan diskretisasi pada $v=87$ ketika menghitung *entropy* dan *gain* pada semua atribut. Hasilnya disajikan pada tabel 3.8.

Tabel 3. 8 Posisi v untuk Pemecahan Atribut *Raport* IPS

<i>Raport</i> IPS	77		82		87		92	
	\leq	$>$	\leq	$>$	\leq	$>$	\leq	$>$
IPA	1	14	7	8	14	1	15	0
IPS	3	12	12	3	15	0	15	0
JUMLAH	4	26	19	11	29	1	30	0
ENTROPY	0,8113	0,9957	0,9495	0,8454	0,9991	0,0000	1,0000	0,0000
GAIN	0,0289		0,0887		0,0342		0,0000	

Hasil penentuan posisi v pada atribut *Raport* IPS menunjukkan nilai *gain* tertinggi didapat pada posisi $v=82$. Maka untuk atribut *Raport* IPS dilakukan diskretisasi pada $v=82$ ketika menghitung *entropy* dan *gain* pada semua atribut. Hasilnya disajikan pada tabel 3.9

Tabel 3. 9 Posisi v untuk Pemecahan Atribut *Raport IQ*

IQ	77		82		87		92	
	\leq	$>$	\leq	$>$	\leq	$>$	\leq	$>$
IPA	0	15	2	13	4	11	8	7
IPS	4	11	6	9	7	8	11	4
JUMLAH	4	26	8	22	11	19	19	11
ENTROPY	0,0000	0,9829	0,8113	0,9760	0,9457	0,9819	0,9819	0,9457
GAIN	0,1482		0,0679		0,0314		0,0314	

Hasil penentuan posisi v pada atribut IQ menunjukkan nilai *gain* tertinggi didapat pada posisi $v = 77$. Maka untuk atribut IQ dilakukan diskretisasi pada $v = 77$ ketika menghitung *entropy* dan *gain* pada semua atribut.

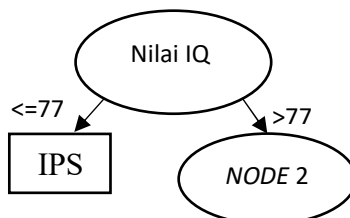
Setelah semua atribut diperoleh *entropy* dan *gain* tertinggi maka dihitung *gain* untuk masing-masing atribut dipilih dengan nilai yang paling besar pula. Hasilnya disajikan pada tabel 3.10.

Tabel 3. 10 Hasil Perhitungan *Entropy* dan *Gain* untuk *Node Akar*

ATRIBUT		IPA	IPS	<i>Entropy</i>	<i>Gain</i>
Nilai <i>Raport B INDO</i>	≤ 77	0	3	0,0000	0,1080
	> 77	15	12	0,9911	
Nilai <i>Raport B INGG</i>	≤ 92	15	14	0,9991	0,0342
	> 92	0	1	0,0000	
Nilai <i>Raport MTK</i>	≤ 82	7	12	0,9495	0,0887
	> 82	8	3	0,8454	
Nilai <i>Raport IPA</i>	≤ 87	13	15	0,9963	0,0701
	> 87	2	0	0,0000	
Nilai <i>Raport IPS</i>	≤ 82	7	12	0,9495	0,0887
	> 82	8	3	0,8454	
Nilai IQ	≤ 77	0	4	0,0000	0,1482
	> 77	15	11	0,9829	
MINAT	IPA	4	11	1,0000	0,0000
	IPS	4	11	1,0000	
Nilai Max					0,1482

Hasil yang didapat di tabel 3.10 menunjukkan bahwa *gain* tertinggi ada di atribut Nilai IQ, maka Nilai IQ dijadikan sebagai *node akar*. Pada atribut Nilai IQ *entropy* yang dihasilkan pada nilai ≤ 77 sehingga atribut

Nilai IQ dijadikan syarat kondisi di *node* 1 dan *node* 2. Pohon yang terbentuk untuk proses pada *node* 2 adalah sebagai berikut



Gambar 3. 2 Pohon Keputusan yang Terbentuk pada *Node* akar

Tabel 3. 11 Data Kasus pada Nilai IQ ≤ 77

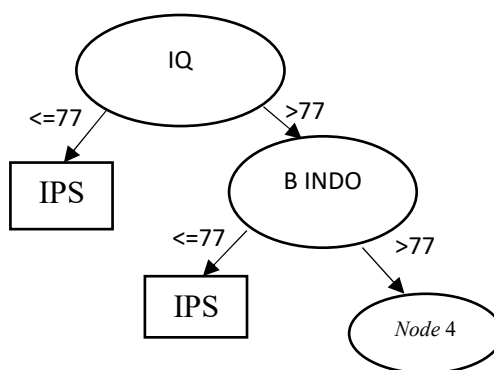
NO	R INDO	R B ING	R MM	R IPA	R IPS	IQ	MINAT	HASIL
1	82	83	78	83	78	94	IPS	IPA
2	88	84	85	84	85	106	IPS	IPA
3	87	90	86	94	86	113	IPA	IPA
4	87	91	87	94	87	94	IPA	IPA
5	86	84	81	81	81	86	IPS	IPA
6	78	79	79	80	79	81	IPA	IPA
7	89	85	84	82	84	91	IPA	IPA
8	84	80	78	80	78	98	IPS	IPA
9	86	90	85	82	85	84	IPS	IPA
10	78	86	76	77	76	81	IPS	IPA
11	89	84	88	83	88	91	IPS	IPA
12	80	82	81	79	81	113	IPS	IPA
13	91	83	83	82	83	91	IPS	IPA
14	87	82	78	81	78	94	IPS	IPA
15	89	85	85	82	85	91	IPS	IPA
16	80	88	76	78	76	94	IPS	IPS
17	77	81	76	77	76	94	IPS	IPS
18	77	78	78	75	78	81	IPS	IPS
19	87	85	82	83	82	89	IPS	IPS
20	87	87	83	87	83	106	IPA	IPS
21	85	79	81	83	81	91	IPS	IPS
22	79	83	77	78	77	84	IPA	IPS
23	88	94	86	84	86	113	IPS	IPS
24	84	84	80	80	80	91	IPS	IPS
25	83	81	83	85	83	91	IPA	IPS
26	87	83	82	84	82	81	IPS	IPS

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk atribut Nilai IQ ≤ 77 . Hasil pemecahan tersebut menyisakan 26 data yang terdiri dari 15 data dengan hasil IPA dan 11 data dengan hasil IPS. Hasil dari perhitungan atribut diperoleh *entropy* dan *gain* tertinggi maka dihitung *gain* untuk masing-masing atribut dipilih dengan nilai yang paling besar pula. Hasilnya disajikan pada tabel 3.12

Tabel 3. 12 Hasil perhitungan *entropy* dan *gain* untuk *node 2*

ATRIBUT		IPA	IPS	<i>Entropy</i>	<i>Gain</i>
Nilai <i>Raport B INDO</i>	≤ 77	0	2	0,0000	0,1018
	> 77	15	9	0,9544	
Nilai <i>Raport B INGG</i>	≤ 92	15	0	0,9710	0,0493
	> 92	10	1	0,0000	
Nilai <i>Raport MTK</i>	≤ 77	1	14	0,8113	0,0579
	> 77	3	8	0,9457	
Nilai <i>Raport IPA</i>	≤ 87	13	2	0,9950	0,0644
	> 87	11	0	0,0000	
Nilai <i>Raport IPS</i>	≤ 77	1	14	0,8113	0,0579
	> 77	3	8	0,9457	
Nilai IQ	≤ 92	8	7	0,9968	0,0077
	> 92	7	4	0,9457	
MINAT	IPA	4	11	0,9852	0,0000
	IPS	3	8	0,9819	
Nilai Max					0,1018

Hasil yang didapat di tabel 3.12 menunjukkan bahwa *gain* tertinggi ada di atribut Nilai *Raport B INDO*, maka Nilai *Raport B INDO* dijadikan sebagai *node* internal (*node 2*). Pada atribut Nilai *Raport B INDO* *entropy* yang dihasilkan pada nilai ≤ 77 memiliki nilai 0 sehingga dapat dipastikan semua data latih yang memiliki Nilai *Raport B INDO* ≤ 77 masuk dalam kelas IPS. Sedangkan nilai *entropy* yang > 77 akan dijadikan sebagai *node* selanjutnya. Pohon yang terbentuk untuk proses pada *node 2* adalah sebagai berikut



Gambar 3. 3 Pohon Keputusan yang Terbentuk pada *Node 2*

Tabel 3. 13 Data Kasus pada Nilai *raport B INDO >77*

NO	R INDO	R B ING	R MM	R IPA	R IPS	IQ	MINAT	HASIL
1	82	83	78	83	78	94	IPS	IPA
2	88	84	85	84	85	106	IPS	IPA
3	87	90	86	94	86	113	IPA	IPA
4	87	91	87	94	87	94	IPA	IPA
5	86	84	81	81	81	86	IPS	IPA
6	78	79	79	80	79	81	IPA	IPA
7	89	85	84	82	84	91	IPA	IPA
8	84	80	78	80	78	98	IPS	IPA
9	86	90	85	82	85	84	IPS	IPA
10	78	86	76	77	76	81	IPS	IPA
11	89	84	88	83	88	91	IPS	IPA
12	80	82	81	79	81	113	IPS	IPA
13	91	83	83	82	83	91	IPS	IPA
14	87	82	78	81	78	94	IPS	IPA
15	89	85	85	82	85	91	IPS	IPA
16	80	88	76	78	76	94	IPS	IPS
17	87	85	82	83	82	89	IPS	IPS
18	87	87	83	87	83	106	IPA	IPS
19	85	79	81	83	81	91	IPS	IPS
20	79	83	77	78	77	84	IPA	IPS
21	88	94	86	84	86	113	IPS	IPS
22	84	84	80	80	80	91	IPS	IPS
23	83	81	83	85	83	91	IPA	IPS
24	87	83	82	84	82	81	IPS	IPS

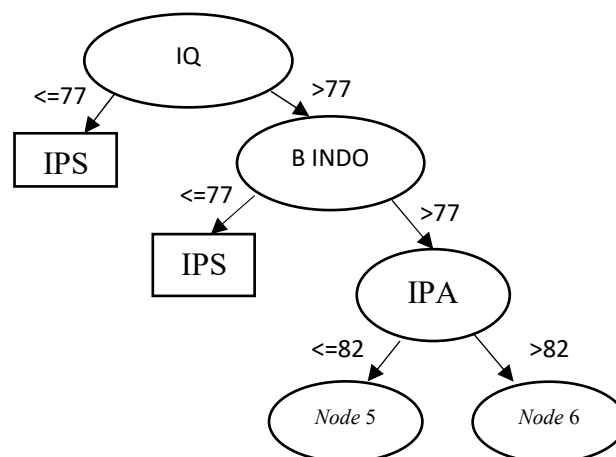
Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk atribut Nilai *Raport B INDO >77*. Hasil pemecahan tersebut menyisakan 24 data yang terdiri dari 15 data dengan hasil IPA dan 9 data dengan hasil IPS. Hasil dari

perhitungan atribut diperoleh *entropy* dan *gain* tertinggi maka dihitung *gain* untuk masing-masing atribut dipilih dengan nilai yang paling besar pula. Hasilnya disajikan pada tabel 3.14

Tabel 3. 14 Hasil perhitungan *entropy* dan *gain* untuk *node 3*

ATRIBUT		IPA	IPS	<i>Entropy</i>	<i>Gain</i>
Nilai Raport B INDO	<= 87	10	8	0,9911	0,0486
	> 87	5	1	0,6500	
Nilai Raport B INGG	<= 92	15	8	0,9321	0,0612
	> 92	0	1	0,0000	
Nilai Raport MTK	<= 77	1	2	0,9183	0,0361
	> 77	14	7	0,9183	
Nilai Raport IPA	<= 82	10	3	0,7793	0,0767
	>82	5	6	0,9940	
Nilai Raport IPS	<= 77	1	2	0,9183	0,0361
	> 77	14	7	0,9183	
Nilai IQ	<= 92	8	6	0,9852	0,0125
	> 92	7	3	0,8813	
MINAT	IPA	4	11	0,9852	0,0036
	IPS	3	6	0,9367	
Nilai Max					0,0767

Hasil yang didapat di tabel 3.14 menunjukkan bahwa *gain* tertinggi ada di atribut Nilai *Raport* IPA, maka Nilai *Raport* IPA dijadikan sebagai *node* internal (*node 4*). Pada atribut Nilai *Raport* IPA tidak digunakan lagi karena kedua nilai berbeda yang tersisa digunakan untuk syarat pengujian *node 5*. Selanjutnya ada *node 6* untuk fitur numerik kembali dilakukan perhitungan posisi v buat pemecahan. Pohon yang terbentuk untuk proses pada *node 4* adalah sebagai berikut.



Gambar 3. 4 Pohon Keputusan yang Terbentuk pada *Node 4*

Tabel 3. 15 Data Kasus pada Nilai *raport* IPA ≤ 82

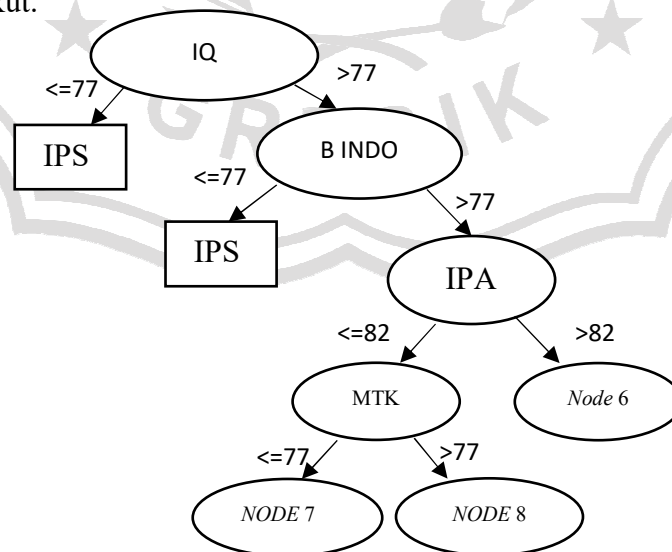
NO	R INDO	R B ING	R MM	R IPA	R IPS	IQ	MINAT	HASIL
1	86	84	81	81	81	86	IPS	IPA
2	78	79	79	80	79	81	IPA	IPA
3	89	85	84	82	84	91	IPA	IPA
4	84	80	78	80	78	98	IPS	IPA
5	86	90	85	82	85	84	IPS	IPA
6	78	86	76	77	76	81	IPS	IPA
7	80	82	81	79	81	113	IPS	IPA
8	91	83	83	82	83	91	IPS	IPA
9	87	82	78	81	78	94	IPS	IPA
10	89	85	85	82	85	91	IPS	IPA
11	80	88	76	78	76	94	IPS	IPS
12	79	83	77	78	77	84	IPA	IPS
13	84	84	80	80	80	91	IPS	IPS

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk atribut Nilai *raport* IPA ≤ 82 . Hasil pemecahan tersebut menyisakan 13 data yang terdiri dari 10 data dengan hasil IPA dan 3 data dengan hasil IPS. Hasil dari perhitungan atribut diperoleh *entropy* dan *gain* tertinggi maka dihitung *gain* untuk masing-masing atribut dipilih dengan nilai yang paling besar pula. Hasilnya disajikan pada tabel 3.16

Tabel 3. 16 Data Kasus pada *Raport* IPA ≤ 82

ATRIBUT		IPA	IPS	Entropy	Gain
Nilai <i>Raport</i> B INDO	≤ 87	7	3	0,8813	0,1014
	> 87	3	0	0,0000	
Nilai <i>Raport</i> B INGG	≤ 87	4	0	0,0000	0,1436
	> 87	6	3	0,9183	
Nilai <i>Raport</i> MTK	≤ 77	1	2	0,9183	0,2067
	> 77	9	1	0,4690	
Nilai <i>Raport</i> IPA	≤ 77	1	0	0,0000	0,0305
	> 77	9	3	0,8113	
Nilai <i>Raport</i> IPS	≤ 77	1	2	0,9183	0,2067
	> 77	9	1	0,4690	
Nilai IQ	≤ 82	2	0	0,0000	0,0641
	> 82	8	3	0,8454	
MINAT	IPA	2	8	0,9183	0,0121
	IPS	1	2	0,7219	
Nilai Max					0,2067

Hasil yang didapat di tabel 3.16 menunjukkan bahwa *gain* tertinggi ada di atribut Nilai *Raport* MTK, maka Nilai *Raport* MTK dijadikan sebagai *node* internal (*node* 5). Pada atribut Nilai *Raport* MTK tidak digunakan lagi karena kedua nilai berbeda yang tersisa digunakan untuk syarat pengujian *node* 7. Selanjutnya ada *node* 8 untuk fitur numerik kembali dilakukan perhitungan posisi v buat pemecahan. Pohon yang terbentuk untuk proses pada *node* 5 adalah sebagai berikut.

Gambar 3. 5 Pohon Keputusan yang Terbentuk pada *Node* 5

Tabel 3. 17 Data Kasus pada Nilai *raport* MTK ≤ 77

NO	R INDO	R B ING	R MM	R IPA	R IPS	IQ	MINAT	HASIL
1	78	86	76	77	76	81	IPS	IPA
2	80	88	76	78	76	94	IPS	IPS
3	79	83	77	78	77	84	IPA	IPS

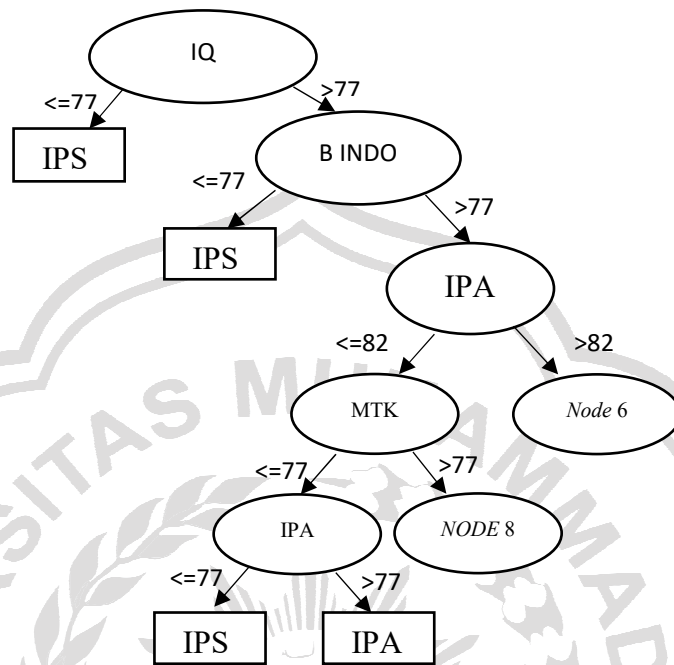
Perhitungan selanjutnya adalah dengan menggunkan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk atribut Nilai *raport* MTK ≤ 77 . Hasil pemecahan tersebut menyisakan 3 data yang terdiri dari 1 data dengan hasil IPA dan 2 data dengan hasil IPS. Hasil dari perhitungan atribut diperoleh *entropy* dan *gain* tertinggi maka dihitung *gain* untuk masing-masing atribut dipilih dengan nilai yang paling besar pula. Hasilnya disajikan pada tabel 3.18

Tabel 3. 18 Data Kasus pada *Raport* MTK ≤ 77

ATRIBUT		IPA	IPS	<i>Entropy</i>	<i>Gain</i>
Nilai <i>Raport</i> B INDO	≤ 77	0	0	0,0000	0,0000
	> 77	1	2	0,9183	
Nilai <i>Raport</i> B INGG	≤ 87	1	1	1,0000	0,2516
	> 87	0	1	0,0000	
Nilai <i>Raport</i> MTK	≤ 77	1	2	0,9183	0,0000
	> 77	0	0	0,0000	
Nilai <i>Raport</i> IPA	≤ 77	1	0	0,0000	0,9183
	> 77	0	2	0,0000	
Nilai <i>Raport</i> IPS	≤ 77	1	2	0,9183	0,0000
	> 77	0	1	0,0000	
Nilai IQ	≤ 82	1	0	0,0000	0,9183
	> 82	0	2	0,0000	
MINAT	IPA	0	1	0,0000	0,2516
	IPS	1	1	1,0000	
Nilai Max					0,9183

Hasil yang didapat di tabel 3.18 menunjukkan bahwa *gain* tertinggi ada di atribut Nilai *Raport* IPA, maka Nilai *Raport* IPA dijadikan sebagai *node* internal (*node* 7). Pada atribut Nilai *Raport* IPA *entropy* yang dihasilkan pada nilai ≤ 77 memiliki nilai 0 sehingga dapat dipastikan semua data latih yang memiliki Nilai *Raport* IPA ≤ 77 masuk dalam kelas

IPS. Sedangkan nilai *entropy* yang Nilai *Raport* IPA >77 masuk dalam kelas IPA. Pohon yang terbentuk untuk proses pada *node 7* adalah sebagai berikut.



Gambar 3. 6 Pohon Keputusan yang Terbentuk pada Node 7

Tabel 3. 19 Data Kasus pada Nilai *raport* MTK > 77

NO	R INDO	R B ING	R MM	R IPA	R IPS	IQ	MINAT	HASIL
1	86	84	81	81	81	86	IPS	IPA
2	78	79	79	80	79	81	IPA	IPA
3	89	85	84	82	84	91	IPA	IPA
4	84	80	78	80	78	98	IPS	IPA
5	86	90	85	82	85	84	IPS	IPA
6	80	82	81	79	81	113	IPS	IPA
7	91	83	83	82	83	91	IPS	IPA
8	87	82	78	81	78	94	IPS	IPA
9	89	85	85	82	85	91	IPS	IPA
10	84	84	80	80	80	91	IPS	IPS

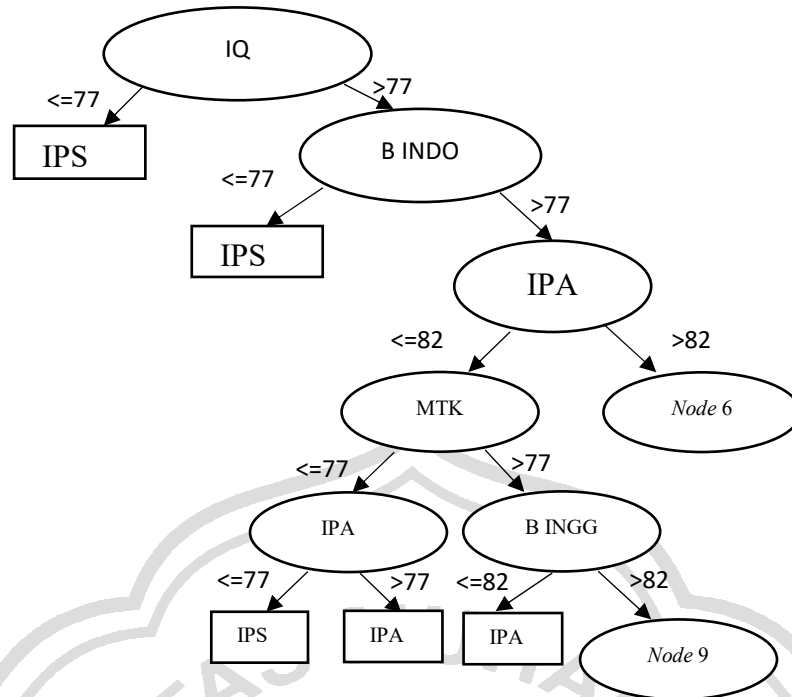
Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk atribut Nilai *raport* MTK >77. Hasil pemecahan tersebut menyisakan 10 data yang terdiri dari 9 data dengan hasil IPA dan 1 data dengan hasil IPS.

Hasil dari perhitungan atribut diperoleh *entropy* dan *gain* tertinggi maka dihitung *gain* untuk masing-masing atribut dipilih dengan nilai yang paling besar pula. Hasilnya disajikan pada tabel 3.20

Tabel 3. 20 Data Kasus pada *Raport* MTK >77

ATRIBUT		IPA	IPS	<i>Entropy</i>	<i>Gain</i>
Nilai <i>Raport</i> B INDO	≤ 87	6	1	0,5917	0,0548
	> 87	3	0	0,0000	
Nilai <i>Raport</i> B INGG	≤ 82	4	0	0,0000	0,0790
	> 82	5	1	0,6500	
Nilai <i>Raport</i> MTK	≤ 82	5	1	0,6500	0,0790
	> 82	4	0	0,0000	
Nilai <i>Raport</i> IPA	≤ 77	0	0	0,0000	0,0000
	> 77	9	1	0,4690	
Nilai <i>Raport</i> IPS	≤ 82	5	1	0,6500	0,0790
	> 82	4	0	0,0000	
Nilai IQ	≤ 87	3	0	0,0000	0,0548
	> 87	6	1	0,5917	
MINAT	IPA	2	7	0,0000	0,0341
	IPS	0	1	0,5436	
Nilai Max					0,0790

Hasil yang didapat di tabel 3.20 menunjukkan bahwa *gain* tertinggi ada di atribut Nilai *Raport* B INGG, maka Nilai *Raport* B INGG dijadikan sebagai *node* internal (*node* 8). Pada atribut Nilai *Raport* B INGG *entropy* yang dihasilkan pada nilai ≤ 82 memiliki nilai 0 sehingga dapat dipastikan semua data latih yang memiliki Nilai *Raport* B INGG ≤ 82 masuk dalam kelas IPA. Sedangkan nilai *entropy* yang Nilai *Raport* B INGG > 82 dijadikan sebagai *node* 9. Pohon yang terbentuk untuk proses pada *node* 8 adalah sebagai berikut.



Gambar 3. 7 Pohon Keputusan yang Terbentuk pada Node 8

Tabel 3. 21 Data Kasus pada Nilai *raport* B INGG > 82

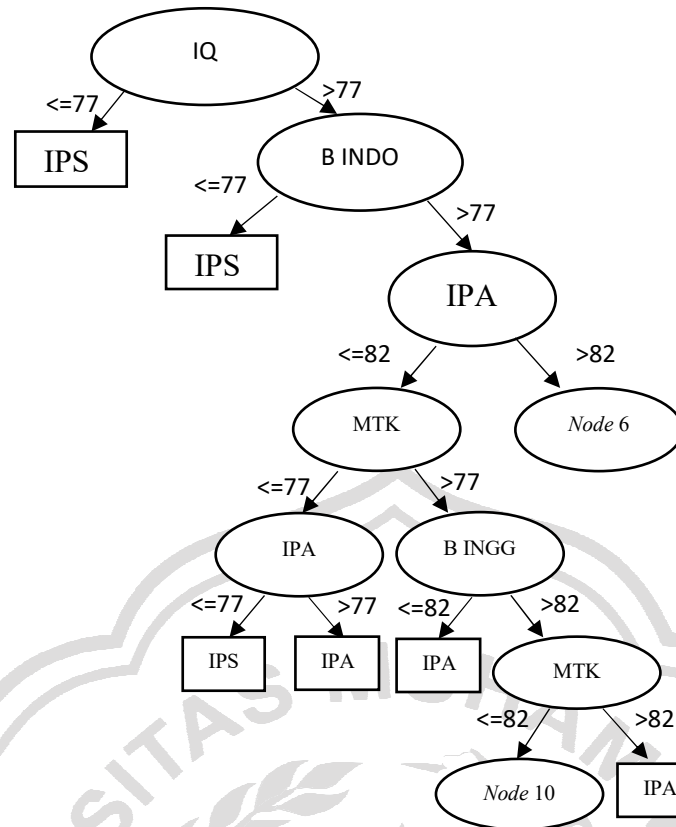
NO	R INDO	R B ING	R MM	R IPA	R IPS	IQ	MINAT	HASIL
1	86	84	81	81	81	86	IPS	IPA
2	89	85	84	82	84	91	IPA	IPA
3	86	90	85	82	85	84	IPS	IPA
4	91	83	83	82	83	91	IPS	IPA
5	89	85	85	82	85	91	IPS	IPA
6	84	84	80	80	80	91	IPS	IPS

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk atribut Nilai *raport* B INGG > 82. Hasil pemecahan tersebut menyisakan 6 data yang terdiri dari 5 data dengan hasil IPA dan 1 data dengan hasil IPS. Hasil dari perhitungan atribut diperoleh *entropy* dan *gain* tertinggi maka dihitung *gain* untuk masing-masing atribut dipilih dengan nilai yang paling besar pula. Hasilnya disajikan pada tabel 3.22

Tabel 3. 22 Data Kasus pada *Raport B INGG >82*

ATRIBUT		IPA	IPS	Entropy	Gain
Nilai <i>Raport B INDO</i>	≤ 87	2	1	0,9183	0,1909
	> 87	3	0	0,0000	
Nilai <i>Raport B INGG</i>	≤ 87	4	1	0,7219	0,0484
	> 87	1	0	0,0000	
Nilai <i>Raport MTK</i>	≤ 82	1	1	1,0000	0,3167
	> 82	4	0	0,0000	
Nilai <i>Raport IPA</i>	≤ 77	0	0	0,0000	0,0000
	> 77	5	1	0,6500	
Nilai <i>Raport IPS</i>	≤ 82	1	1	1,0000	0,3167
	> 82	4	0	0,0000	
Nilai IQ	≤ 87	2	0	0,0000	0,1092
	> 87	3	1	0,8113	
MINAT	IPA	1	4	0,0000	0,0484
	IPS	0	1	0,7219	
Nilai Max					0,3167

Hasil yang didapat di tabel 3.23 menunjukkan bahwa *gain* tertinggi ada di atribut Nilai *Raport MTK*, maka Nilai *Raport MTK* dijadikan sebagai *node* internal (*node* 9). Pada atribut Nilai *Raport MTK* *entropy* yang dihasilkan pada nilai >82 memiliki nilai 0 sehingga dapat dipastikan semua data latih yang memiliki Nilai *Raport MTK* >82 masuk dalam kelas IPA. Sedangkan nilai *entropy* yang Nilai *Raport MTK* ≤ 82 dijadikan sebagai *node* 10. Pohon yang terbentuk untuk proses pada *node* 9 adalah sebagai berikut.



Gambar 3. 8 Pohon Keputusan yang Terbentuk pada *Node 9*

Tabel 3. 23 Data Kasus pada Nilai *raport* MTK ≤ 82

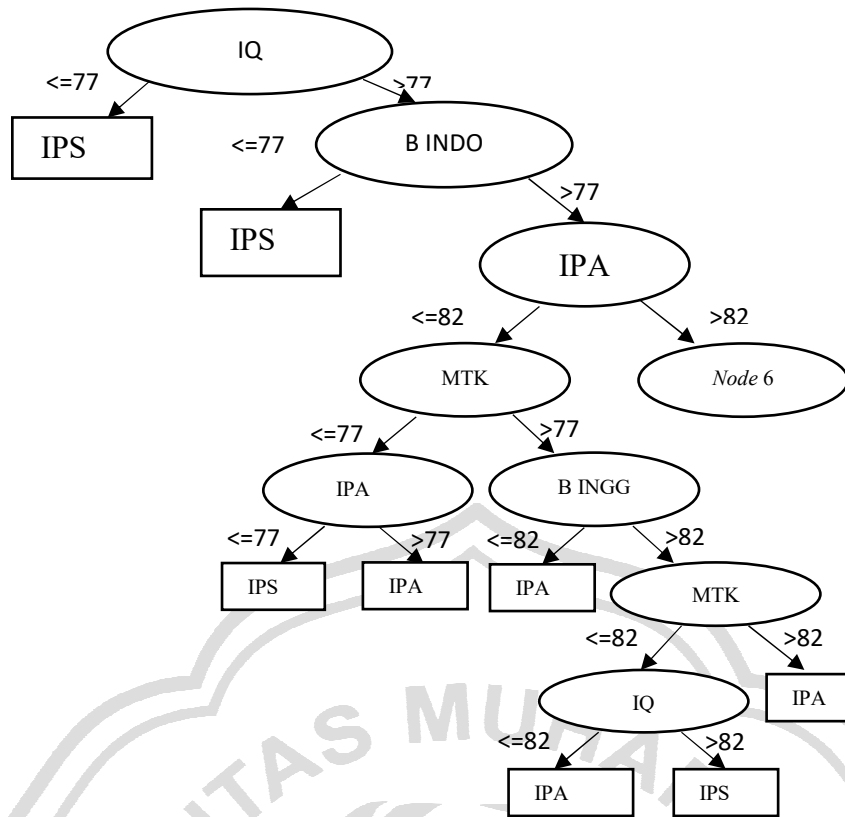
NO	R INDO	R B ING	R MM	R IPA	R IPS	IQ	MINAT	HASIL
1	86	84	81	81	81	86	IPS	IPA
2	84	84	80	80	80	91	IPS	IPS
2	84	84	80	80	80	91	IPS	IPS

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk atribut Nilai *raport* MTK ≤ 82 . Hasil pemecahan tersebut menyisakan 2 data yang terdiri dari 1 data dengan hasil IPA dan 1 data dengan hasil IPS. Hasil dari perhitungan atribut diperoleh *entropy* dan *gain* tertinggi maka dihitung *gain* untuk masing-masing atribut dipilih dengan nilai yang paling besar pula. Hasilnya disajikan pada tabel 3.24

Tabel 3. 24 Data Kasus pada *Raport MTK* ≤ 82

ATRIBUT		IPA	IPS	<i>Entropy</i>	<i>Gain</i>
Nilai <i>Raport B INDO</i>	≤ 77	0	0	0,0000	0,0000
	> 77	1	1	1,0000	
Nilai <i>Raport B INGG</i>	≤ 77	0	0	0,0000	0,0000
	> 77	1	1	1,0000	
Nilai <i>Raport MTK</i>	≤ 77	0	0	0,0000	0,0000
	> 77	1	1	1,0000	
Nilai <i>Raport IPA</i>	≤ 77	0	0	0,0000	0,0000
	> 77	1	1	1,0000	
Nilai <i>Raport IPS</i>	≤ 77	0	0	0,0000	0,0000
	> 77	1	1	1,0000	
Nilai IQ	≤ 87	1	0	0,0000	1,0000
	> 87	0	1	0,0000	
MINAT	IPA	0	1	0,0000	0,0000
	IPS	0	1	1,0000	
Nilai Max					1,0000

Hasil yang didapat di tabel 3.24 menunjukkan bahwa *gain* tertinggi ada di atribut Nilai IQ, maka Nilai IQ dijadikan sebagai *node* internal (*node* 10). Pada atribut Nilai IQ *entropy* yang dihasilkan pada nilai ≤ 87 memiliki nilai 0 sehingga dapat dipastikan semua data latih yang memiliki Nilai IQ ≤ 87 masuk dalam kelas IPA. Sedangkan nilai *entropy* yang Nilai IQ > 87 masuk dalam kelas IPS. Pohon yang terbentuk untuk proses pada *node* 10 adalah sebagai berikut.



Gambar 3. 9 Pohon Keputusan yang Terbentuk pada Node 10

Tabel 3. 25 Data Kasus pada Nilai *Raport* IPA >82

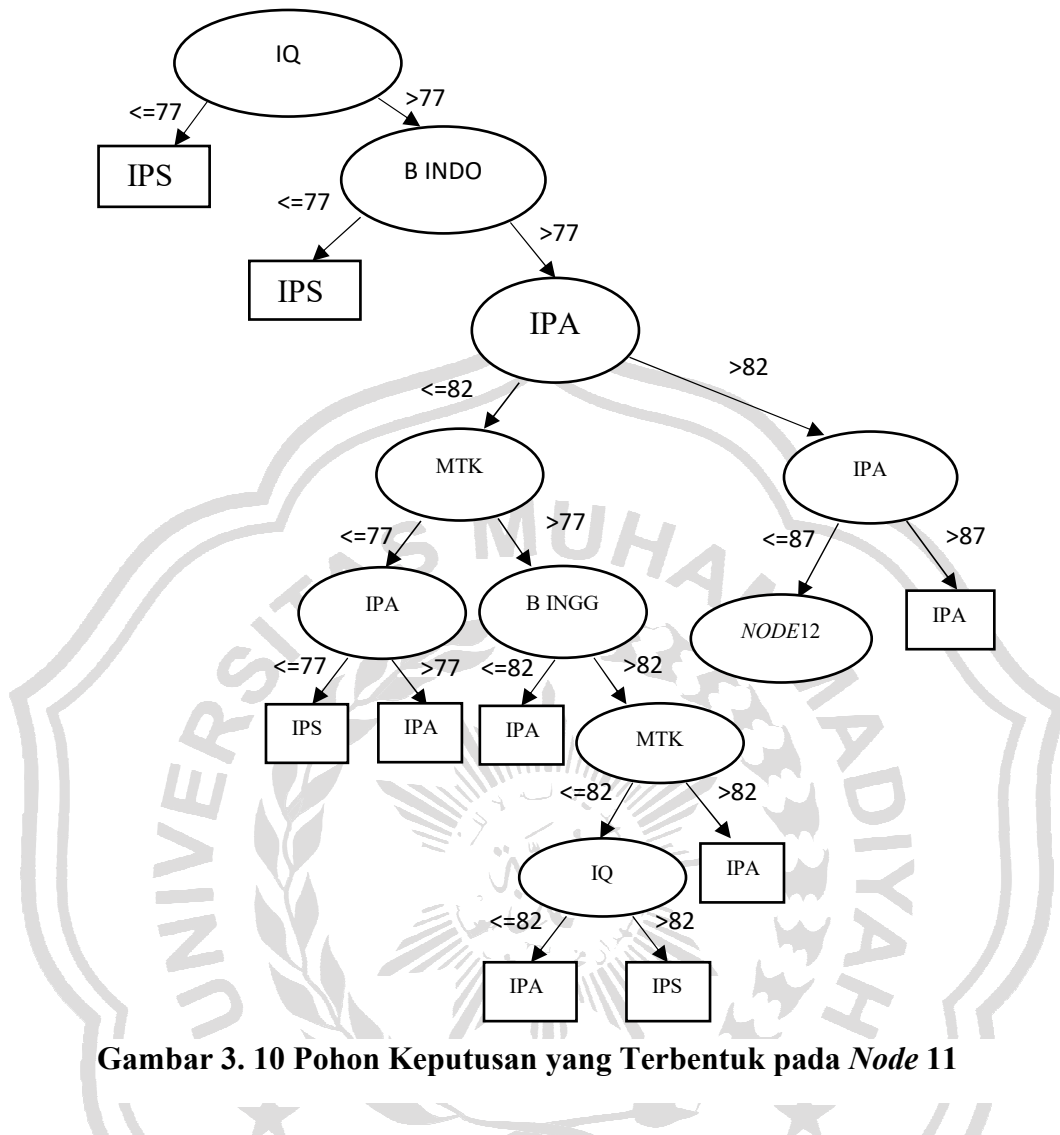
NO	R INDO	R B ING	R MM	R IPA	R IPS	IQ	MINAT	HASIL
1	82	83	78	83	78	94	IPS	IPA
2	88	84	85	84	85	106	IPS	IPA
3	87	90	86	94	86	113	IPA	IPA
4	87	91	87	94	87	94	IPA	IPA
5	89	84	88	83	88	91	IPS	IPA
6	87	85	82	83	82	89	IPS	IPS
7	87	87	83	87	83	106	IPA	IPS
8	85	79	81	83	81	91	IPS	IPS
9	88	94	86	84	86	113	IPS	IPS
10	83	81	83	85	83	91	IPA	IPS
11	87	83	82	84	82	81	IPS	IPS

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk atribut Nilai *Raport* IPA >82. Hasil pemecahan tersebut menyisakan 11 data yang terdiri dari 5 data dengan hasil IPA dan 6 data dengan hasil IPS. Hasil dari perhitungan atribut diperoleh *entropy* dan *gain* tertinggi maka dihitung *gain* untuk masing-masing atribut dipilih dengan nilai yang paling besar pula. Hasilnya disajikan pada tabel 3.26

Tabel 3. 26 Data Kasus pada Nilai *Raport* IPA >82

ATRIBUT		IPA	IPS	<i>Entropy</i>	<i>Gain</i>
Nilai <i>Raport</i> B INDO	<= 82	0	2	0,0000	0,1831
	> 82	5	4	0,9911	
Nilai <i>Raport</i> B INGG	<= 87	4	6	0,9710	0,1113
	> 87	1	0	0,0000	
Nilai <i>Raport</i> MTK	<= 77	0	0	0,0000	0,0000
	> 77	1	1	1,0000	
Nilai <i>Raport</i> IPA	<= 87	3	6	0,9183	0,2427
	>87	2	0	0,0000	
Nilai <i>Raport</i> IPS	<= 87	4	6	0,9710	0,1113
	> 87	1	0	0,0000	
Nilai IQ	<= 92	1	4	0,7219	0,1650
	> 92	4	2	0,9183	
MINAT	IPA	0	1	0,0000	0,0000
	IPS	0	1	1,0000	
Nilai Max					1,0000

Hasil yang didapat di tabel 3.26 menunjukkan bahwa *gain* tertinggi ada di atribut Nilai *Raport* IPA, maka Nilai *Raport* IPA dijadikan sebagai *node* internal (*node* 11). Pada atribut Nilai *Raport* IPA *entropy* yang dihasilkan pada nilai >87 memiliki nilai 0 sehingga dapat dipastikan semua data latih yang memiliki Nilai *Raport* IPA >87 masuk dalam kelas IPA. Sedangkan nilai *entropy* yang Nilai *Raport* IPA <=82 dijadikan sebagai *node* 11. Pohon yang terbentuk untuk proses pada *node* 6 adalah sebagai berikut.



Gambar 3. 10 Pohon Keputusan yang Terbentuk pada *Node 11*

Tabel 3. 27 Data Kasus pada Nilai *Raport IPA <=87*

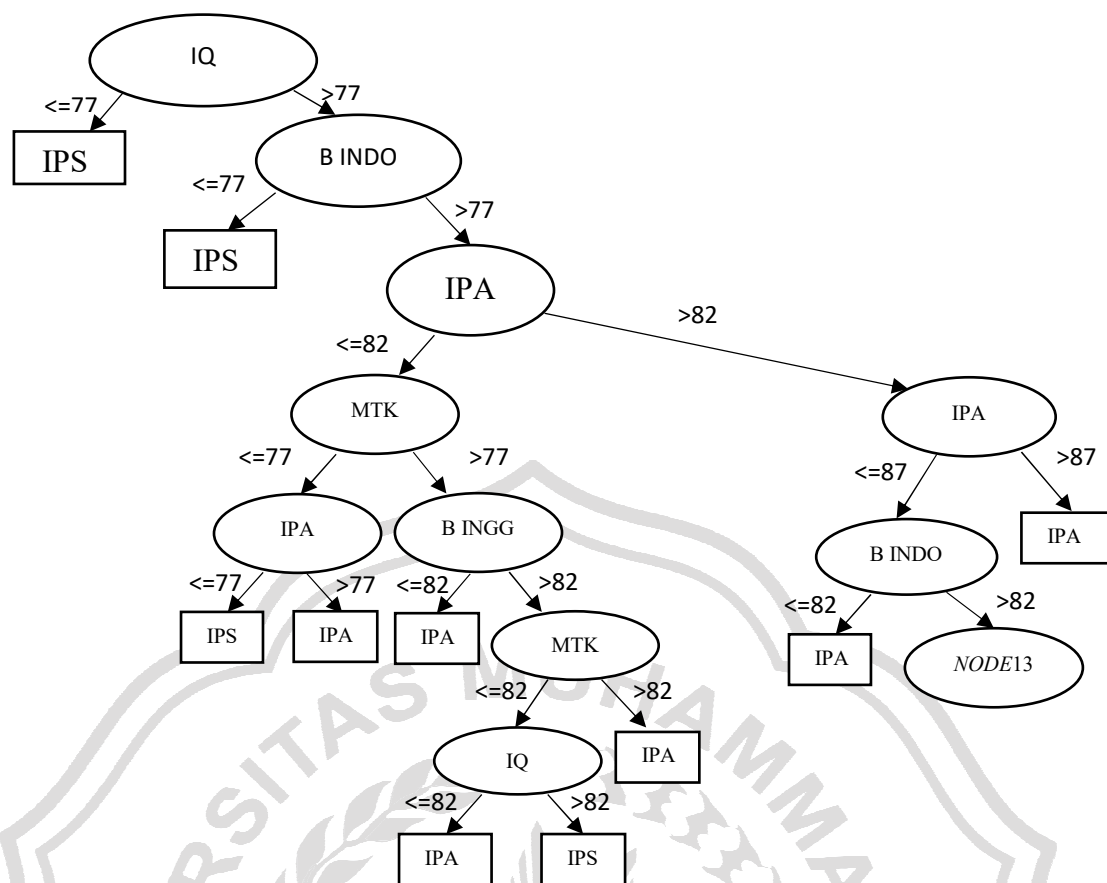
NO	R INDO	R B ING	R MM	R IPA	R IPS	IQ	MINAT	HASIL
1	82	83	78	83	78	94	IPS	IPA
2	88	84	85	84	85	106	IPS	IPA
3	89	84	88	83	88	91	IPS	IPA
4	87	85	82	83	82	89	IPS	IPS
5	87	87	83	87	83	106	IPA	IPS
6	85	79	81	83	81	91	IPS	IPS
7	88	94	86	84	86	113	IPS	IPS
8	83	81	83	85	83	91	IPA	IPS
9	87	83	82	84	82	81	IPS	IPS

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk atribut Nilai *Raport* IPA ≤ 87 . Hasil pemecahan tersebut menyisakan 9 data yang terdiri dari 3 data dengan hasil IPA dan 6 data dengan hasil IPS. Hasil dari perhitungan atribut diperoleh *entropy* dan *gain* tertinggi maka dihitung *gain* untuk masing-masing atribut dipilih dengan nilai yang paling besar pula. Hasilnya disajikan pada tabel 3.28

Tabel 3. 28 Data Kasus pada Nilai *Raport* \leq IPA 87

ATRIBUT		IPA	IPS	<i>Entropy</i>	<i>Gain</i>
Nilai <i>Raport</i> B INDO	≤ 82	1	0	0,0000	0,1972
	> 82	2	6	0,8113	
Nilai <i>Raport</i> B INGG	≤ 82	0	2	0,0000	0,1520
	> 82	3	4	0,9852	
Nilai <i>Raport</i> MTK	≤ 87	2	6	0,8113	0,1972
	> 87	1	0	0,0000	
Nilai <i>Raport</i> IPA	≤ 77	0	0	0,0000	0,0000
	> 77	3	6	0,9183	
Nilai <i>Raport</i> IPS	≤ 87	2	6	0,8113	0,1972
	> 87	1	0	0,0000	
Nilai IQ	≤ 92	1	4	0,7219	0,0728
	> 92	2	2	1,0000	
MINAT	IPA	0	3	0,0000	0,1520
	IPS	2	4	0,9852	
Nilai Max					0,1972

Hasil yang didapat di tabel 3.28 menunjukkan bahwa *gain* tertinggi ada di atribut Nilai *Raport* B INDO, maka Nilai *Raport* B INDO dijadikan sebagai *node* internal (*node* 12). Pada atribut Nilai *Raport* B INDO *entropy* yang dihasilkan pada nilai ≤ 82 memiliki nilai 0 sehingga dapat dipastikan semua data latih yang memiliki Nilai *Raport* B INDO ≤ 82 masuk dalam kelas IPA. Sedangkan nilai *entropy* yang Nilai *Raport* B INDO > 82 dijadikan sebagai *node* 13. Pohon yang terbentuk untuk proses pada *node* 12 adalah sebagai berikut.



Gambar 3. 11 Pohon Keputusan yang Terbentuk pada *Node 12*

Tabel 3. 29 Data Kasus pada Nilai *Raport B INDO >82*

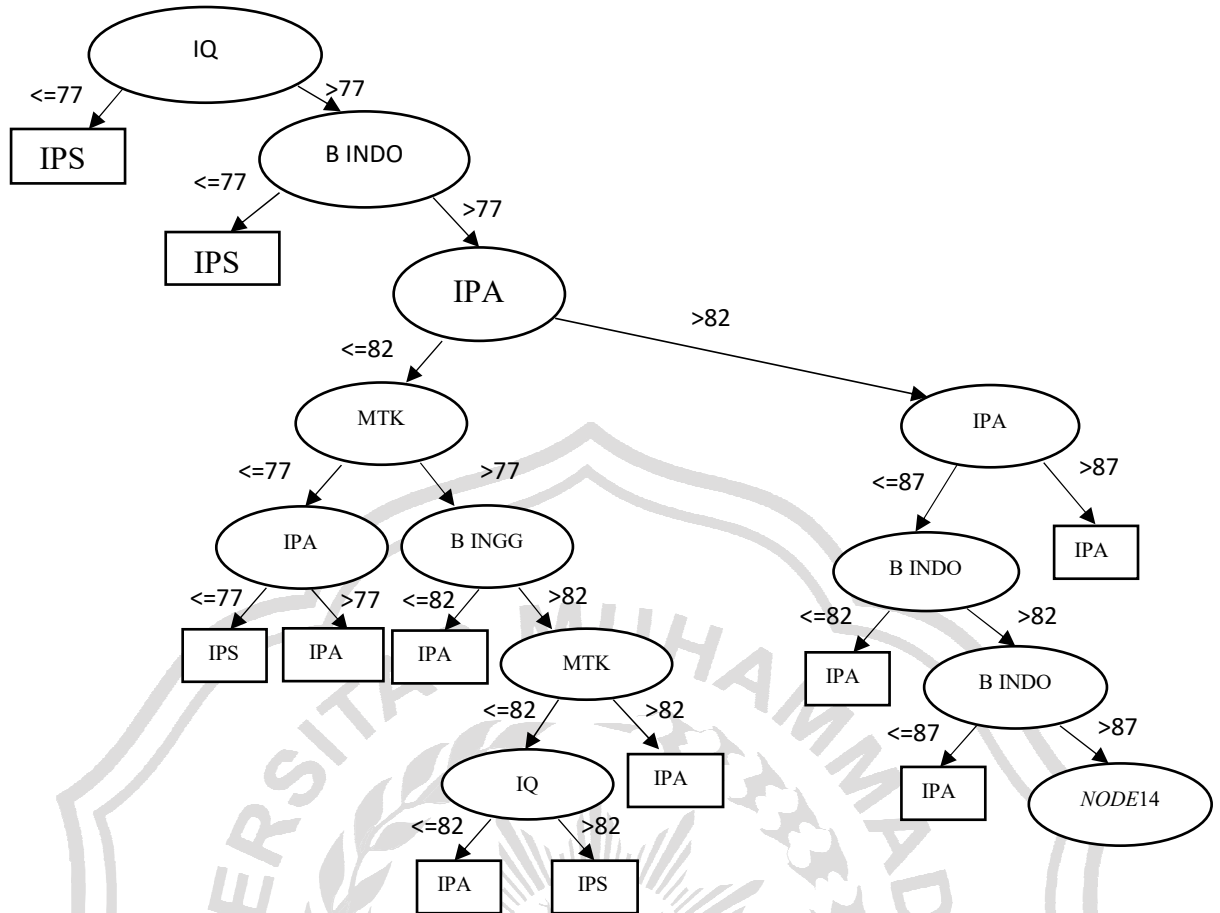
NO	R INDO	R B ING	R MM	R IPA	R IPS	IQ	MINAT	HASIL
1	88	84	85	84	85	106	IPS	IPA
2	89	84	88	83	88	91	IPS	IPA
3	87	85	82	83	82	89	IPS	IPS
4	87	87	83	87	83	106	IPA	IPS
5	85	79	81	83	81	91	IPS	IPS
6	88	94	86	84	86	113	IPS	IPS
7	83	81	83	85	83	91	IPA	IPS
8	87	83	82	84	82	81	IPS	IPS

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk atribut Nilai *Raport B INDO >82*. Hasil pemecahan tersebut menyisakan 8 data yang terdiri dari 2 data dengan hasil IPA dan 6 data dengan hasil IPS. Hasil dari perhitungan atribut diperoleh *entropy* dan *gain* tertinggi maka dihitung *gain* untuk masing-masing atribut dipilih dengan nilai yang paling besar pula. Hasilnya disajikan pada tabel 3.30

Tabel 3. 30 Data Kasus pada Nilai *Raport B INDO* >82

ATRIBUT		IPA	IPS	<i>Entropy</i>	<i>Gain</i>
Nilai <i>Raport B INDO</i>	≤ 87	0	5	0,0000	0,4669
	> 87	2	1	0,9183	
Nilai <i>Raport B INGG</i>	≤ 82	0	2	0,0000	0,1226
	> 82	2	4	0,9183	
Nilai <i>Raport MTK</i>	≤ 87	1	6	0,5917	0,2936
	> 87	1	0	0,0000	
Nilai <i>Raport IPA</i>	≤ 77	0	0	0,0000	0,0000
	> 77	2	6	0,8113	
Nilai <i>Raport IPS</i>	≤ 87	1	6	0,5917	0,2936
	> 87	1	0	0,0000	
Nilai IQ	≤ 87	1	0	0,0000	1,0000
	> 87	0	1	0,0000	
MINAT	IPA	0	2	0,0000	0,1226
	IPS	2	4	0,9183	
Nilai Max					0,4669

Hasil yang didapat di tabel 3.30 menunjukkan bahwa *gain* tertinggi ada di atribut Nilai *Raport B INDO*, maka Nilai *Raport B INDO* dijadikan sebagai *node* internal (*node* 13). Pada atribut Nilai *Raport B INDO* *entropy* yang dihasilkan pada nilai ≤ 87 memiliki nilai 0 sehingga dapat dipastikan semua data latih yang memiliki Nilai *Raport B INDO* ≤ 87 masuk dalam kelas IPS. Sedangkan nilai *entropy* yang Nilai *Raport B INDO* > 87 dijadikan sebagai *node* 14. Pohon yang terbentuk untuk proses pada *node* 13 adalah sebagai berikut.



Gambar 3. 12 Pohon Keputusan yang Terbentuk pada Node 13

Tabel 3. 31 Data Kasus pada Nilai *Raport B INDO* >87

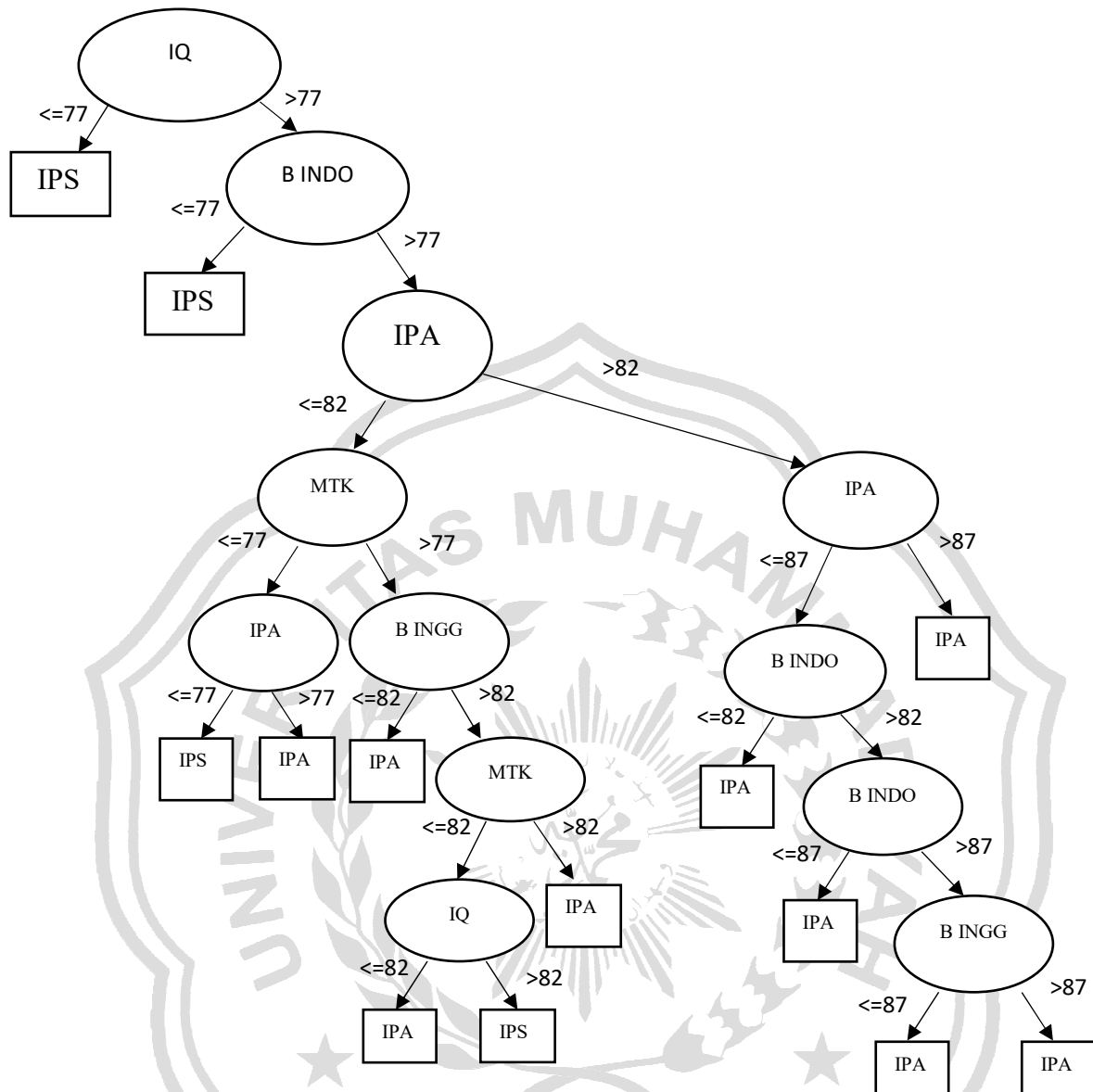
NO	R INDO	R B ING	R MM	R IPA	R IPS	IQ	MINAT	HASIL
1	88	84	85	84	85	106	IPS	IPA
2	89	84	88	83	88	91	IPS	IPA
3	88	94	86	84	86	113	IPS	IPS

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk atribut Nilai *Raport B INDO* >87. Hasil pemecahan tersebut menyisakan 3 data yang terdiri dari 2 data dengan hasil IPA dan 1 data dengan hasil IPS. Hasil dari perhitungan atribut diperoleh *entropy* dan *gain* tertinggi maka dihitung *gain* untuk masing-masing atribut dipilih dengan nilai yang paling besar pula. Hasilnya disajikan pada tabel 3.32

Tabel 3. 32 Data Kasus pada Nilai *Raport B INDO* >87

ATRIBUT		IPA	IPS	<i>Entropy</i>	<i>Gain</i>
Nilai <i>Raport B INDO</i>	≤ 77	0	2	0,0000	0,0000
	> 77	0	1	0,9183	
Nilai <i>Raport B INGG</i>	≤ 87	2	0	0,0000	0,9183
	> 87	0	1	0,0000	
Nilai <i>Raport MTK</i>	≤ 87	1	1	1,0000	0,2516
	> 87	1	0	0,0000	
Nilai <i>Raport IPA</i>	≤ 77	0	2	0,0000	0,0000
	> 77	0	1	0,9183	
Nilai <i>Raport IPS</i>	≤ 87	1	1	1,0000	0,2516
	> 87	1	0	0,0000	
Nilai IQ	≤ 92	1	1	1,0000	0,2516
	> 92	0	1	0,0000	
MINAT	IPA	0	2	0,0000	0,0000
	IPS	0	1	0,9183	
Nilai Max					0,9183

Hasil yang didapat di tabel 3.32 menunjukkan bahwa *gain* tertinggi ada di atribut Nilai B INGG, maka Nilai B INGG dijadikan sebagai *node* internal (*node* 14). Pada atribut Nilai IQ *entropy* yang dihasilkan pada nilai ≤ 87 memiliki nilai 0 sehingga dapat dipastikan semua data latih yang memiliki Nilai B INGG ≤ 87 masuk dalam kelas IPA. Sedangkan nilai *entropy* yang Nilai B INGG > 87 masuk dalam kelas IPS. Pohon yang terbentuk untuk proses pada *node* 14 adalah sebagai berikut



Gambar 3. 13 Pohon Keputusan yang Terbentuk pada *Node 14*

Dari pohon keputusan tersebut akan dijadikan dalam bentuk aturan *IF THEN* sebagai berikut :

RULE 1 = IF (IQ<=77) THEN jurusan IPS

RULE 2 = IF (IQ>77) AND (B.INDONESIA<=77) THEN jurusan IPS

RULE 3 = IF (IQ >77) AND (B.INDONESIA>77) AND (IPA<=82) AND (MATEMATIKA<=77) AND (IPA<=77) THEN jurusan IPA

RULE 4 = IF (IQ >77) AND (B.INDONESIA>77) AND (IPA<=82) AND (MATEMATIKA<=77) AND (IPA>77) THEN jurusan IPS

RULE 5 = IF (IQ >77) AND (B.INDONESIA>77) AND (IPA<=82) AND (MATEMATIKA>77) AND (B.INGGRIS<=82) THEN jurusan IPA

RULE 6 = IF (IQ >77) AND (B.INDONESIA>77) AND (IPA<=82) AND (MATEMATIKA>77) AND (B.INGGRIS>82) AND (MATEMATIKA<=82) AND (IQ<=87) THEN jurusan IPA

RULE 7 = IF (IQ >77) AND (B.INDONESIA>77) AND (IPA<=82) AND (MATEMATIKA>77) AND (B.INGGRIS>82) AND (MATEMATIKA<=82) AND (IQ>87) THEN jurusan IPS

RULE 8 = IF (IQ >77) AND (B.INDONESIA>77) AND (IPA<=82) AND (MATEMATIKA>77) AND (B.INGGRIS>82) AND (MATEMATIKA>82) THEN jurusan IPA

RULE 9 = IF (IQ>77) AND (B.INDONESIA>77) AND (IPA>82) AND (IPA<=87) AND (B.INDONESIA<=82) THEN jurusan IPA

RULE 10 = IF (IQ>77) AND (B.INDONESIA>77) AND (IPA>82) AND (IPA<=87) AND (B.INDONESIA>82) AND (B.INDONESIA<=87) THEN jurusan IPS

RULE 11 = IF (IQ>77) AND (B.INDONESIA>77) AND (IPA>82) AND (IPA<=87) AND (B.INDONESIA>82) AND (B.INDONESIA>87) AND (B.INGGRIS<=87) THEN jurusan IPA

RULE 12 = IF (IQ>77) AND (B.INDONESIA>77) AND (IPA>82) AND (IPA<=87) AND (B.INDONESIA>82) AND (B.INDONESIA>87) AND (B.INGGRIS>87) THEN jurusan IPS

RULE 13 = IF (IQ>77) AND (B.INDONESIA>77) AND (IPA>82) AND (IPA>87) THEN jurusan IPA

3.6 Kebutuhan Pembuatan Sistem

3.1.6 Kebutuhan Perangkat Keras

Perangkat keras adalah alat yang digunakan untuk menunjang dalam pembuatan sistem. Dalam pembuatan sistem ini perangkat keras yang digunakan yaitu laptop dengan spesifikasi :

- Processor Intel Core i3
- RAM 2 GB
- HDD 500 GB
- Monitor 14"

3.1.7 Kebutuhan Perangkat Lunak

Perangkat lunak adalah program atau aplikasi yang digunakan untuk membangun sistem. Perangkat lunak yang dibutuhkan dalam pembuatan sistem ini adalah :

- Windows 8.1 pro
- Web Server : Apache
- Database Server : MySQL
- Bahasa Pemrograman : PHP
- *Sublime Text*
- SQLyog
- *Web Browser* : Google Chrome

3.7 Perancangan Sistem

Bagian ini akan menjelaskan rancangan sistem seperti *context diagram*, diagram berjenjang dan *data flow diagram* (DFD)

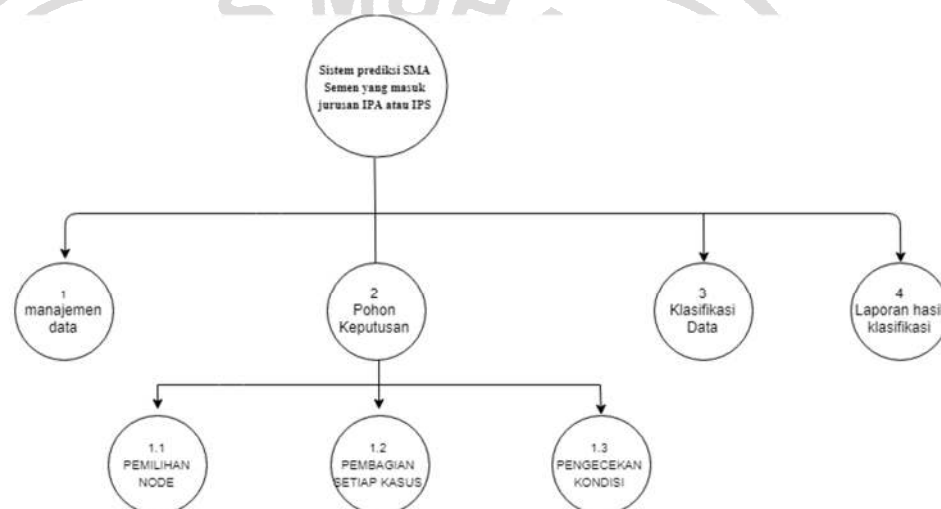
3.1.1 Context Diagram



Gambar 3. 14 Context Diagram

Penjelasan dari gambar 3.14, terlihat bahwa yang terlibat (*entity*) dalam sistem ini adalah Admin (Guru BK) dan *User* (Kepala sekolah). Admin, memasukkan data siswa berupa nilai *raport* nilai IQ dan minat kelas 10, serta kelas prediksi. Data tersebut digunakan sebagai *data training* atau data yang akan diproses untuk pembentukan pohon keputusan. Keluaran dari sistem untuk *user* adalah hasil uji prediksi siswa berupa kelas siswa masuk kelas IPA atau kelas IPS berdasarkan data yang telah dimasukkan. Sedangkan Admin dapat melihat laporan atau daftar hasil uji prediksi siswa yang masuk kelas IPA atau IPS

3.1.2 Diagram Berjenjang



Gambar 3. 15 Diagram Berjenjang

Diagram berjenjang disajikan pada gambar 3.15. berikut penjelasannya:

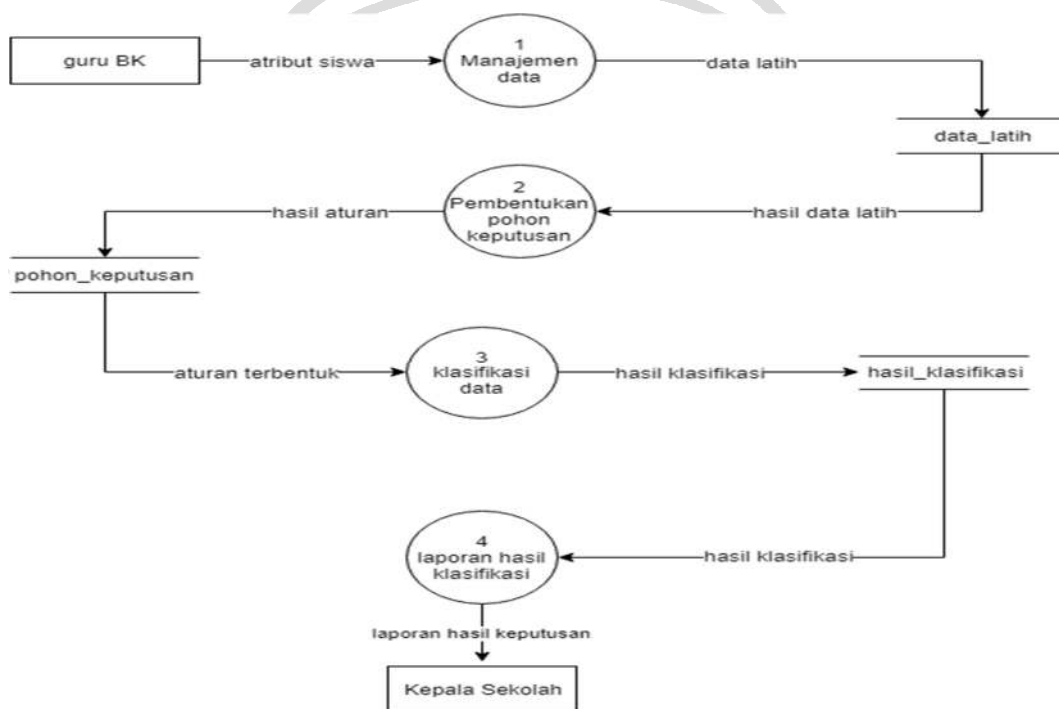
1. *Top level* : Penerapan Metode *Decision Tree* C4.5 untuk menentukan jurusan SMA Semen Gresik.
2. *Level 0* :
 - 1 Manajemen data, merupakan proses pengolahan data *training* dan data uji yang akan digunakan dalam pembentukan pohon keputusan.
 - 2 Pembentukan aturan (pohon keputusan) dengan metode C4.5, yang didalamnya terdapat tiga proses.

3 Pengklasifikasian data uji menggunakan aturan yang sudah terbentuk.

4 Pembuatan laporan hasil klasifikasi

3. *Level 1* :
- 2.1 Pemilihan *node* sebagai pemecah cabang.
 - 2.2 Pembagian cabang pada setiap kasus.
 - 2.3 Pengecekan kondisi, yaitu jika masih ada kasus yang memiliki kelas yang berbeda maka mengulangi.

3.1.3 DFD Level 0



Gambar 3. 16 DFD Level 0 Sistem Penjurusan Siswa

DFD *level 0* pada gambar 3.16 menjelaskan aliran data pada sistem. Terdapat empat proses didalam sistem tersebut. Proses satu adalah manajemen data yang diinputkan oleh guru BK. Data atribut siswa akan menjadi data latihan untuk roses pembentukan pohon keputusan. Proses dua adalah pembentukan aturan (pohon keputusan) yang akan digunakan dalam proses pengkalsifikasian data uji. Proses empat adalah pembuatan laporan hasil kalsifikasi siswa dapat masuk jurusan IPA atau IPS yang akan diberikan ke kepala sekolah.

3.1.4 DFD Level 1



Gambar 3. 17 DFD Level 1 Proses Pembentukan (Pohon Keputusan)

DFD level 1 pada gambar 3.17 Proses pembentukan aturan menggunakan metode *Decision Tree* c4.5 ini memiliki tiga proses didalamnya yaitu, proses pemilihan *node* yang akan dijadikan sebagai pemecah cabang, membagi cabang pada setiap kasus, dan proses pengecekan kondisi. Jika ada kasus yang memiliki kelas berbeda, maka akan mengulangi pada proses pemilihan *node*. Hasil dari proses ini adalah aturan atau pohon keputusan yang akan disimpan pada database.

3.8 Struktur Tabel

Struktur tabel ini menjelaskan tabel atau tempat penyimpanan data yang digunakan untuk keperluan sistem yang akan dibangun. Berikut adalah struktur dari tabel-tabel yang akan digunakan.

a. data_latih

Tabel ini digunakan untuk menyimpan data latih atau data yang akan diproses pada pembentukan pohon keputusan. Tabel ini berisi nilai *Raport* dan nilai *try out*. Struktur dari tabel ini dapat dilihat pada tabel 3.33

Tabel 3. 33 Struktur tabel data latih

No	Field_name	Type	Length	Key
1	Id_latih	Int	11	Primary key
2	Rap_indo	Int	11	
3	Rap_inggris	Int	11	
4	Rap_mtk	Int	11	
5	Rap_ipa	Int	11	
6	Rap_ips	Int	11	
7	IQ	Int	11	
8	Minat	Char	1	

b. data_uji

Tabel ini digunakan untuk menyimpan data pengujian, yaitu untuk menguji tingkat akurasi dari pohon keputusan yang terbentuk.

Tabel 3. 34 Struktur tabel data uji

No	Field_name	Type	Length	Key
1	Id_uji	Int	11	Primary key
2	Rap_indo	Int	11	
3	Rap_inggris	Int	11	
4	Rap_mtk	Int	11	
5	Rap_ipa	Int	11	
6	Rap_ips	Int	11	
7	IQ	Int	11	
9	Minat	Char	1	
10	Hasil	Char	1	

c. hasil_klasifikasi

Tabel ini digunakan untuk menyimpan data hasil klasifikasi. Strukturnya sama dengan tabel data latih namun *field* diagnosa diganti dengan hasil.

Tabel 3. 35 Struktur tabel hasil klasifikasi

No	Field_name	Type	Length	Key
1	Id_klasifikasi	Int	11	Primary key
2	Rap_indo	Int	11	
3	Rap_inggris	Int	11	
4	Rap_mtk	Int	11	
5	Rap_ipa	Int	11	
6	Rap_ips	Int	11	
7	IQ	Int	11	
9	Hasil	Char	1	

d. gain

Tabel ini merupakan *temporary* digunakan untuk menampung hasil perhitungan *gain* seperti pada tabel 3.36.

Tabel 3. 36 Struktur tabel gain

No	Name_field	Type	Length	Key
1	Id	Int	11	Primary key
2	<i>node_id</i>	Int	11	
3	Atribut	Varchar	40	
4	<i>Gain</i>	Double		

e. t_user

Tabel *user* ini dibuat untuk secara khusus agar bisa mengakses aplikasi ini. Data dari *user* tersebut tersimpan dalam tabel *user*. Struktur dari tabel *user* dapat dilihat pada tabel 3.37

Tabel 3. 37 Struktur t_user

No	Name_field	Type	Length	Key
1	user_id	Int	25	Primary key
2	Nama	Varchar	50	
3	Username	Varchar	30	
4	Password	Text		
5	Type	Char	1	

f. t_keputusan

Tabel ini menampung hasil dari proses pembentukan pohon keputusan, yaitu menampung aturan-aturan yang telah terbentuk.

Tabel 3. 38 Struktur t_keputusan

No	Field_name	Type	Length	Key
1	Id	Int	11	Primary key
2	Parent	Text		
3	Akar	Text		
4	Keputusan	Varchar	10	

3.9 Desain Antarmuka

Tampilan antar muka pengguna sebagai admin yaitu halaman *login*, *home*, olah data, *mining*, pohon keputusan, hasil klasifikasi, ubah *password* dan *logout*. Sedangkan pengguna sebagai pengguna adalah *login*, *home*, ubah *password* dan laporan hasil klasifikasi.

a) Halaman *Login* (admin maupun user)

Halaman *login* diperlukan untuk mengetahui hak akses pengguna yang masuk kedalam sistem yaitu pengguna sebagai admin.

Halaman Login

Username	<input style="width: 95%;" type="text"/>
Password	<input style="width: 95%;" type="password"/>

Gambar 3. 18 Rancangan halaman login

b) Halaman Utama

Menu yang ditampilkan untuk pengguna sebagai admin yaitu kelola admin, Data Latih, Data Uji dan Klasifikasi C4.5.

Admin	Log Out
Sistem Penentuan Jurusan SMA SEMEN GRESIK	
Kelola Admin Tambah Lihat Data Latih Data Uji Klasifikasi C4 5 Proses Mining Pohon Keputusan Uji Rule Klasifikasi	<div style="text-align: center;">Dashboard</div> <div style="display: flex; justify-content: space-around; margin-top: 10px;"> <div style="border: 1px solid black; padding: 5px; width: 40%;">Data Latih</div> <div style="border: 1px solid black; padding: 5px; width: 40%;">Pohon Keputusan</div> </div>

Gambar 3. 19 Rancangan halaman utama

c) Halaman Data Latih

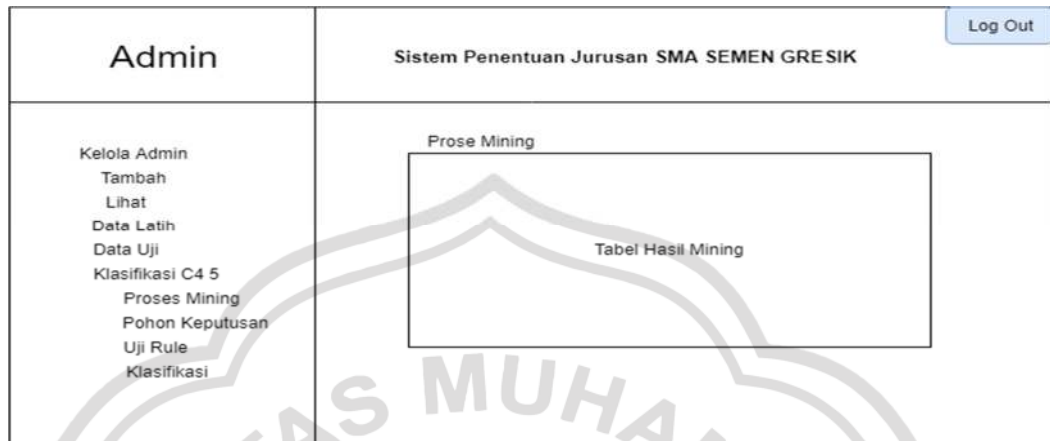
Tampilan ini adalah halaman yang akan digunakan untuk proses pembentukan pohon keputusan

Admin	Log Out										
Sistem Penentuan Jurusan SMA SEMEN GRESIK											
Kelola Admin Tambah Lihat Data Latih Data Uji Klasifikasi C4 5 Proses Mining Pohon Keputusan Uji Rule Klasifikasi	<div style="display: flex; justify-content: space-around; margin-bottom: 10px;"> <div style="border: 1px solid black; padding: 5px; width: 30%;">Tambah Data</div> <div style="border: 1px solid black; padding: 5px; width: 30%;">Hapus Seluruh Data</div> <div style="border: 1px solid black; padding: 5px; width: 30%;">Submit</div> </div> <table border="1" style="width: 100%; border-collapse: collapse; text-align: center;"> <thead> <tr> <th style="width: 5%;">NO</th> <th style="width: 40%;">Raport</th> <th style="width: 10%;">IQ</th> <th style="width: 20%;">Minat</th> <th style="width: 25%;">Action</th> </tr> </thead> <tbody> <tr> <td></td> <td>indo mtk igg IPS IPA 87 86 85 89 96</td> <td>150</td> <td>IPA</td> <td></td> </tr> </tbody> </table>	NO	Raport	IQ	Minat	Action		indo mtk igg IPS IPA 87 86 85 89 96	150	IPA	
NO	Raport	IQ	Minat	Action							
	indo mtk igg IPS IPA 87 86 85 89 96	150	IPA								

Gambar 3. 20 Rancangan halaman data latih

d) Halaman Klasifikasi

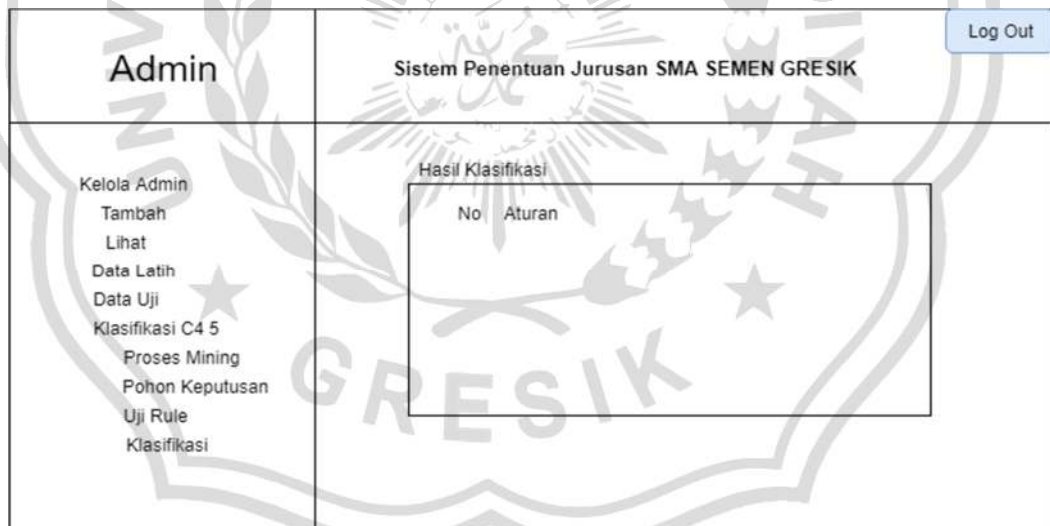
Tampilan ini adalah halaman yang akan digunakan untuk proses pembentukan pohon keputusan dan perhitungan *mining*.



Gambar 3. 21 Rancangan halaman klasifikasi

e) Halaman Hasil Klasifikasi

Halaman ini menampilkan pohon keputusan atau aturan yang didapat dari proses *mining*.



Gambar 3. 22 Rancangan halaman hasil klasifikasi

f) Halaman Uji Akurasi

Halaman ini digunakan untuk menguji tingkat akurasi pohon keputusan yang terbentuk dari proses *mining*.

Admin	Sistem Penentuan Jurusan SMA SEMEN GRESIK Log Out
Kelola Admin Tambah Lihat Data Latih Data Uji Klasifikasi C4 5 Proses Mining Pohon Keputusan Uji Rule Klasifikasi	<div style="text-align: center;"> <p>Hasil Akurasi</p> <div style="border: 1px solid black; padding: 5px; width: fit-content; margin: 0 auto;">Tabel Uji</div> <p>Akurasi...%</p> <p>Laju Error...%</p> <div style="display: flex; justify-content: space-around; margin-top: 10px;"> <div style="border: 1px solid black; padding: 5px; width: 150px; text-align: center;">Gambar Grafik Akurasi dan Laju Error</div> <div style="border: 1px solid black; padding: 5px; width: 150px; text-align: center;">Gambar Grafik Prediksi Jurusan</div> </div> </div>

Gambar 3. 23 Rancangan halaman uji akurasi

g) Halaman Prediksi

Halaman hasil ini akan menampilkan form yang digunakan untuk menambahkan data penggunaan listrik baru yang akan dilakukan prediksi oleh sistem.

Admin	Sistem Penentuan Jurusan SMA SEMEN GRESIK Log Out
Kelola Admin Tambah Lihat Data Latih Data Uji Klasifikasi C4 5 Proses Mining Pohon Keputusan Uji Rule Klasifikasi	<div style="text-align: center;"> <p>Form Data Latih</p> <p>Raport Indo <input type="text"/></p> <p>Raport Inggris <input type="text"/></p> <p>Raport MTK <input type="text"/></p> <p>Raport IPA <input type="text"/></p> <p>Raport IPS <input type="text"/></p> <p>IQ <input type="text"/></p> <p>Minat <input type="text"/></p> <div style="display: flex; justify-content: flex-end; margin-top: 10px;"> Reset Proses </div> </div>

Gambar 3. 24 Rancangan form penambahan prediksi

h) Halaman Tambah Administrator

Halaman ini digunakan untuk menambahkan user/pengguna dalam sistem

Admin	Sistem Penentuan Jurusan SMA SEMEN GRESIK Log Out	
Kelola Admin Tambah Lihat Data Latih Data Uji Klasifikasi C4 5 Proses Mining Pohon Keputusan Uji Rule Klasifikasi	Nama Lengkap <input type="text"/> Email <input type="text"/>	Username <input type="text"/> Password <input type="text"/> <div style="text-align: right;"> <input type="button" value="Reset"/> <input type="button" value="Submit"/> </div>

Gambar 3. 25 Rancangan form tambah administrator

3.10 Evaluasi Sistem

Evaluasi yang digunakan yaitu menggunakan *Confusion Matrik* yaitu tabel yang digunakan untuk menentukan kinerja suatu model klasifikasi. Untuk mengukur nilai akurasi yang didapat dari hasil pengujian, menggunakan rumus 3.1. Sedangkan untuk mengukur tingkat kesalahannya menggunakan rumus 3.2.

$$\text{Akurasi} = \frac{\text{Jumlah } h \text{ data yang diklasifikasi secara benar}}{\text{Jumlah } h \text{ klasifikasi yang dilakukan}} \times 100\% \dots \dots \dots (3.1)$$

$$\text{Laju Error} = \frac{\text{Jumlah } h \text{ data yang diklasifikasi secara salah}}{\text{Jumlah } h \text{ klasifikasi yang dilakukan}} \times 100\% \dots \dots \dots (3.2)$$

Selain itu evaluasi yang digunakan juga dapat dilihat dengan menghitung nilai *Recall*, *Precision*, dan *F-Measure* sehingga dapat dilihat nilai akurasi total dan nilai laju *error*.

3.11 Skenario Pengujian Sistem

Sebelum membuat aplikasi klasifikasi untuk memprediksi siswa yang masuk kelas IPA atau IPS dengan metode *Decision Tree C45* ini, perlu dilakukan beberapa skenario pengujian sistem terlebih dahulu, agar sistem dapat berjalan sesuai dengan tujuan pembuatannya.

- a) Data yang digunakan untuk pengujian sistem terdapat 71 data yang diambil dari periode 2018-2019, dari 71 data tersebut dibagi menjadi 2 macam data yaitu data latih sebanyak 30 dan data uji sebanyak 41. Data latih di gunakan untuk membentuk pohon keputusan dan data uji digunakan untuk menguji akurasi. Selanjutnya membentuk pohon keputusan dari data latih yang sudah di sediakan. Untuk data uji, diklasifikasi berdasarkan pohon keputusan yang terbentuk. Kemudian dihitung akurasi yang menunjukkan baik atau tidaknya pohon keputusan yang sudah terbentuk.
- b) Pada uji hasil prediksi maka sistem dapat menentukan prediksi jurusan siswa yang dapat masuk kelas “IPA” dan “IPS”.
- c) Akurasi sistem diperoleh dari hasil prosentase akurasi ketepatan sistem dan hasil prosentase akurasi kesalahan sistem.