

## BAB III

### ANALISIS DAN PERANCANGAN SISTEM

#### 3.1 Analisis Sistem

Analisis Sistem adalah penguraian suatu sistem yang utuh ke dalam bagian-bagian komponennya yang bertujuan untuk mengidentifikasi dan menganalisa permasalahan, kesempatan, hambatan yang terjadi dan kebutuhan yang diharapkan sehingga dapat diusulkan perbaikannya. Berdasarkan pengamatan, siswa MI Nurul Falah menginginkan masuk ke SMP Negeri, namun sulitnya lolos seleksi dikarenakan banyaknya jumlah pendaftar dan kurangnya Lembaga Pendidikan SMP Negeri yang berada diwilayah Gresik. Kriteria yang dijadikan pedoman untuk menentukan siswa dapat diterima di SMP Negeri adalah berdasarkan Lampiran Keputusan Kepala Dinas Pendidikan Kabupaten Gresik Nomor 421.3/ 555 /437.53/2016 sesuai dengan aspek atau komponen yang diterapkan untuk penilaian . Maka dari itu untuk meningkatkan siswa yang dapat diterima di SMP Negeri diperlukan sebuah sistem yang baik untuk mengetahui prestasi belajar yang diperoleh siswa. Dalam hal ini digunakan sistem prediksi siswa untuk dapat diterima di SMP Negeri. Kriteria yang dijadikan pedoman untuk memprediksi siswa tersebut meliputi nilai Ujian Nasional untuk 3 (tiga mata pelajaran) yaitu Bahasa Indoneisa, Matematika, dan Ilmu Pengetahuan Alam, nilai *try out* dan prestasi olimpiade akademik yang diperoleh siswa.

Setiap atribut atau aspek yang diproses dengan menggunakan metode *Decision Tree C4.5* sebagai sistem yang mampu memprediksi siswa untuk diterima di SMP Negeri dimana hasilnya untuk membantu guru dalam mengetahui siswa yang mengalami kesulitan dalam belajarnya dan untuk siswa dapat memicu untuk belajar lebih giat lagi.

## 3.2 Hasil Analisis

Hasil analisis yang didapat adalah dengan menggunakan sistem ini guru bisa mendapatkan informasi kategori prediksi awal untuk masuk SMP Negeri berupa masuk negeri dan masuk swasta. Pembuatan aplikasi dengan memanfaatkan metode klasifikasi data mining dengan Decision Tree C4.5 memerlukan data latih, data tersebut diperoleh dari nilai Ujian Nasional untuk 3 (tiga) mata pelajaran, nilai *try out* dan prestasi olimpiade akademik dengan prediksi status oleh guru, yang nantinya akan diolah dengan menggunakan metode Decision Tree C4.5. Setiap atribut yang diproses dengan menggunakan metode *Decision Tree C4.5* sebagai sistem yang mampu memprediksi siswa untuk diterima di SMP Negeri dimana hasilnya untuk membantu guru dalam mengetahui siswa yang mengalami kesulitan dalam belajarnya dan untuk siswa dapat memacu untuk belajar lebih giat lagi.

Sistem yang dibangun merupakan aplikasi atau *tool* klasifikasi siswa masuk SMP Negeri menggunakan teknik data mining dengan metode *Decision Tree C4.5*. Sistem ini akan menghasilkan nilai keluaran berupa status prediksi siswa yang masuk kelas negeri dan siswa yang masuk kelas swasta. Terdapat beberapa atribut yang dibutuhkan untuk mengklasifikasi prediksi siswa yang masuk dalam negeri diantaranya: Nilai Ujian Nasional untuk mata pelajaran (Bahasa Indonesia, Matematika, Ilmu Pengetahuan Alam), nilai *try out* dan Prestasi Olimpiade Akademik.

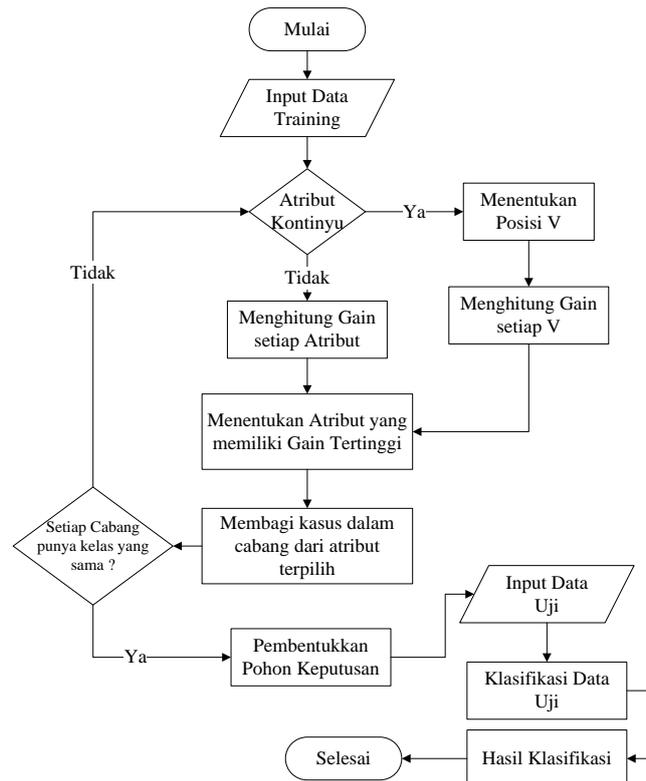
### 3.2.1 Deskripsi Sistem

Sistem pengklasifikasian yang dibangun menggunakan metode *Decision Tree C4.5*. Sistem ini menghasilkan nilai keluaran berupa status prediksi siswa yang masuk kelas negeri dan siswa yang masuk kelas swasta.

Terdapat beberapa variabel yang dibutuhkan untuk mengetahui status prediksi siswa yang masuk dalam SMP Negeri diantaranya adalah Nilai Ujian Nasional untuk mata pelajaran (Bahasa Indonesia,

Matematika, Ilmu Pengetahuan Alam), nilai *try out* dan Prestasi Olimpiade Akademik.

Gambar 3.1 akan menjelaskan alur sistem prediksi siswa yang dapat masuk SMP Negeri menggunakan metode *Decision Tree C4.5*. Algoritma C4.5 merujuk pada gambar 2.2.



**Gambar 3.1** Flowchart System

Penjelasan gambar 3.1:

1. Pertama memasukkan data training (data pasien yang telah diperiksa) yang akan disimpan didalam *database*.
2. Lalu sistem akan memeriksa apakah atribut dari data training bertipe kontinyu (numerik)?
3. Jika atribut dari data training bertipe kontinyu maka sistem akan menentukan posisi V, setelah itu akan dihitung gain setiap posisi V.
4. Jika atribut bertipe kategorikal (bukan kontinyu) maka akan dilakukan penghitungan gain untuk atribut tersebut.

5. Dari hasil penghitungan gain bertipe kontinyu maupun kategorikal tersebut selanjutnya akan dilakukan penentuan cabang/atribut yang memiliki nilai gain paling tinggi.
6. Pembagian kasus dalam cabang dari atribut terpilih.
7. Sistem akan memeriksa apakah setiap cabang memiliki kelas yang sama?
8. Jika setiap cabang memiliki kelas yang berbeda maka dilakukan penghitungan kembali ke poin no.2.
9. Jika setiap cabang memiliki kelas yang sama maka menghasilkan pembentukan pohon keputusan.
10. Selanjutnya masukkan data uji yang akan dilakaukan klasifikasi.
11. Sistem akan melakukan klsifikasi data uji dengan menggunakan pohon keputusan yang sudah terbentuk.
12. Sistem menghasilkan output hasil klasifikasi.

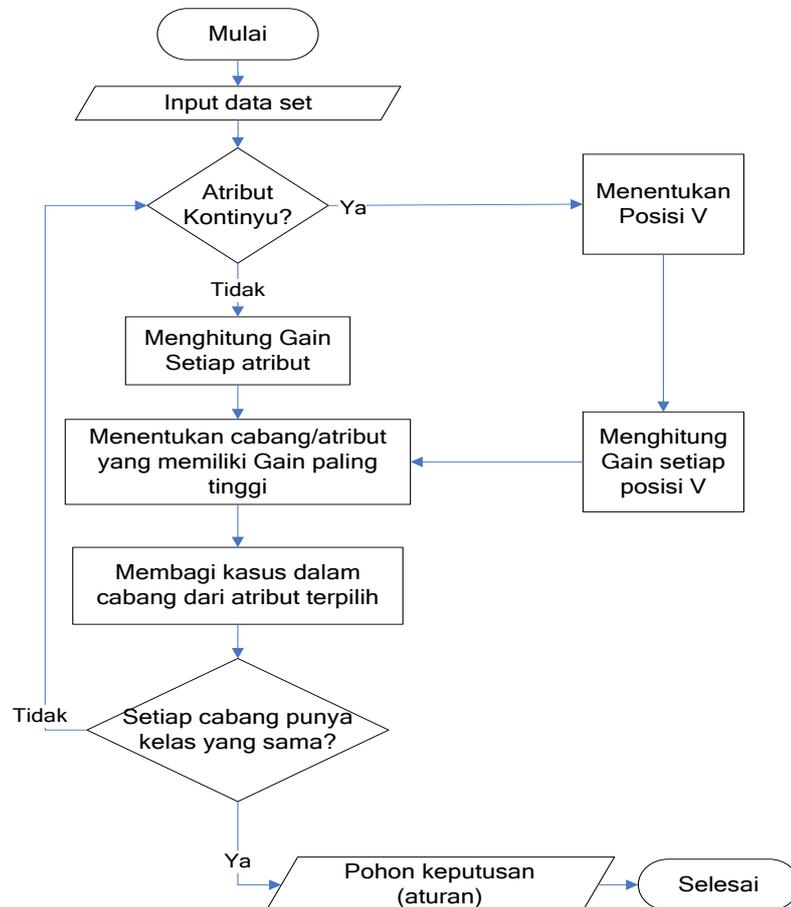
### 3.2.2 Kebutuhan Data

Tahapan awal yang dilakukan dalam penelitian ini adalah menyaipkan data, dimana data diperoleh dari data 4 tahun sebelum dilakukan penelitian. Data yang digunakan adalah data kelulusan 2013-2014 sampai dengan 2016-2017. Data yang diperoleh akan digunakan dalam penelitian ini berupa Nilai Ujian Nasional untuk mata pelajaran (Bahasa Indonesia, Matematika, Ilmu Pengetahuan Alam), nilai *try out* dan Prestasi Olimpiade Akademik. Jumlah data yang digunakan sebanyak 142 record dengan kelas “negeri” dan “swasta” masing-masing berjumlah 68 dan 74 yang akan dibagi menjadi data latih dan data uji.

### 3.2.3 Proses Perhitungan Metode *Decision Tree C4.5*

Metode pengklasifikasian dengan menggunakan *Decisison Tree C4.5* adalah pengembangan dari *ID3*, jika pada *ID3* pengklasifikasian hanya bisa dilakukan untuk atribut yang memiliki nilai kategorikal saja,

maka *C4.5* dapat dilakukan untuk atribut yang memiliki nilai kategorika maupun nilai kontinyu (numerik). Berikut ini adalah alur diagram dari algoritma *Decision Tree C4.5* digambarkan pada Gambar 3.2



**Gambar 3.2** Algoritma metode *Decision Tree C4.5*

Untuk memilih atribut sebagai simpul akar (*root node*) atau simpul dalam (*internal node*), didasarkan pada nilai *information gain* tertinggi dari atribut-atribut yang ada. Sebelum perhitungan *information gain*, akan dilakukan perhitungan *entropy*. *Entropy* merupakan distribusi probabilitas dalam teori informasi dan diadopsi kedalam algoritma *C4.5* untuk mengukur tingkat homogenitas distribusi kelas dari sebuah himpunan data (*data set*). Semakin tinggi tingkat *entropy* dari sebuah data maka semakin homogen distribusi kelas pada data tersebut.

### 3.3 Representasi Data

Dari 142 data siswa, diambil 120 data yang akan dijadikan data latih dan 22 sebagai data uji. Data latih disajikan pada tabel 3.1. Sedangkan data uji disajikan pada tabel 3.2.

**Tabel 3.1** Data latih

NO	NILAI RAPOR			NILAI TRY OUT			AKADMIK	SEKOLAH
	IPA	BINDO	MTK	IPA	BINDO	MTK		
1	85	78	81	82	78	80	TIDAK ADA	SWASTA
2	73	78	70	74	78	71	TIDAK ADA	SWASTA
3	73	78	70	73	78	70	TIDAK ADA	SWASTA
4	74	78	70	73	78	70	TIDAK ADA	SWASTA
5	78	78	70	74	78	70	TIDAK ADA	SWASTA
6	73	78	70	73	78	71	TIDAK ADA	SWASTA
7	73	78	70	73	78	70	TIDAK ADA	SWASTA
8	73	78	70	74	78	71	TIDAK ADA	SWASTA
9	80	78	70	73	78	71	TIDAK ADA	SWASTA
10	73	78	70	74	78	70	TIDAK ADA	SWASTA
11	86	78	74	76	79	72	TIDAK ADA	SWASTA
12	79	78	71	84	81	78	TIDAK ADA	SWASTA
13	84	78	70	74	79	71	TIDAK ADA	SWASTA
14	73	78	70	73	78	70	TIDAK ADA	SWASTA
15	73	78	70	75	78	71	TIDAK ADA	SWASTA
16	75	78	77	74	78	77	TIDAK ADA	SWASTA
17	74	79	70	73	78	73	TIDAK ADA	SWASTA
18	79	79	76	78	79	76	TIDAK ADA	SWASTA
19	73	78	70	74	78	70	TIDAK ADA	SWASTA
20	73	78	70	73	78	70	TIDAK ADA	SWASTA
21	73	78	70	73	78	70	TIDAK ADA	SWASTA
22	76	78	71	75	83	78	TIDAK ADA	SWASTA
23	78	78	72	78	78	74	TIDAK ADA	SWASTA
24	73	78	70	75	80	73	TIDAK ADA	SWASTA
25	75	78	71	74	79	71	TIDAK ADA	SWASTA
26	83	81	75	78	88	75	TIDAK ADA	SWASTA
27	74	78	70	74	90	72	TIDAK ADA	SWASTA
28	76	83	70	89	95	80	TIDAK ADA	SWASTA
29	76	81	76	80	86	80	TIDAK ADA	SWASTA
30	78	78	70	89	92	72	TIDAK ADA	SWASTA
31	78	78	72	77	87	73	TIDAK ADA	SWASTA
32	76	79	77	83	84	80	TIDAK ADA	SWASTA

33	80	78	92	75	78	85	TIDAK ADA	SWASTA
34	76	78	70	74	78	70	TIDAK ADA	SWASTA
35	73	78	91	74	78	75	TIDAK ADA	SWASTA
36	76	78	70	74	78	70	TIDAK ADA	SWASTA
37	83	85	70	80	87	73	TIDAK ADA	SWASTA
38	74	78	70	74	81	82	TIDAK ADA	SWASTA
39	75	78	73	74	85	82	TIDAK ADA	SWASTA
40	76	83	73	85	84	82	TIDAK ADA	SWASTA
41	74	78	70	75	78	70	TIDAK ADA	SWASTA
42	76	85	70	75	81	70	TIDAK ADA	SWASTA
43	83	85	78	87	88	78	TIDAK ADA	SWASTA
44	74	78	72	75	79	72	TIDAK ADA	SWASTA
45	73	78	70	74	78	70	TIDAK ADA	SWASTA
46	76	80	71	76	79	70	TIDAK ADA	SWASTA
47	74	82	70	78	84	76	TIDAK ADA	SWASTA
48	73	78	70	76	79	70	TIDAK ADA	SWASTA
49	74	78	70	73	78	70	TIDAK ADA	SWASTA
50	73	80	70	73	78	70	TIDAK ADA	SWASTA
51	74	78	80	74	81	73	TIDAK ADA	SWASTA
52	73	78	70	73	78	70	TIDAK ADA	SWASTA
53	85	78	70	80	84	70	TIDAK ADA	SWASTA
54	73	78	70	73	80	70	TIDAK ADA	SWASTA
55	77	80	73	77	81	73	TIDAK ADA	SWASTA
56	73	83	74	83	87	75	TIDAK ADA	SWASTA
57	73	78	70	73	78	70	TIDAK ADA	SWASTA
58	73	78	71	75	80	75	TIDAK ADA	SWASTA
59	73	78	71	73	78	70	TIDAK ADA	SWASTA
60	81	86	72	79	84	77	TIDAK ADA	SWASTA
75	91	84	82	84	84	85	TIDAK ADA	NEGERI
76	83	84	84	87	86	85	TIDAK ADA	NEGERI
77	91	78	80	87	88	95	TIDAK ADA	NEGERI
78	94	78	89	83	85	91	TIDAK ADA	NEGERI
79	94	83	83	81	84	85	TIDAK ADA	NEGERI
80	90	81	76	89	87	91	TIDAK ADA	NEGERI
81	92	84	85	88	89	93	TIDAK ADA	NEGERI
82	86	83	83	84	84	83	TIDAK ADA	NEGERI
83	90	82	81	80	82	85	TIDAK ADA	NEGERI
84	83	83	81	87	84	84	TIDAK ADA	NEGERI
85	87	83	85	82	86	85	TIDAK ADA	NEGERI
86	89	81	91	90	86	97	TIDAK ADA	NEGERI
87	82	83	82	82	85	85	TIDAK ADA	NEGERI

88	85	95	83	83	78	83	TIDAK ADA	NEGERI
89	83	84	83	83	82	81	TIDAK ADA	NEGERI
90	81	83	82	82	86	85	TIDAK ADA	NEGERI
91	87	85	82	82	86	85	TIDAK ADA	NEGERI
92	83	84	82	83	84	83	TIDAK ADA	NEGERI
93	85	83	80	83	83	83	TIDAK ADA	NEGERI
94	86	90	83	87	90	90	TIDAK ADA	NEGERI
95	84	84	83	85	87	93	TIDAK ADA	NEGERI
96	84	81	83	85	84	83	TIDAK ADA	NEGERI
97	83	83	82	83	83	82	TIDAK ADA	NEGERI
98	85	86	83	84	84	83	TIDAK ADA	NEGERI
99	78	84	80	87	88	83	TIDAK ADA	NEGERI
100	84	92	82	84	92	82	TIDAK ADA	NEGERI
101	89	93	78	93	92	86	ADA	NEGERI
102	83	83	82	84	83	82	TIDAK ADA	NEGERI
103	83	78	84	84	78	84	TIDAK ADA	NEGERI
104	81	84	83	81	83	83	TIDAK ADA	NEGERI
105	80	82	83	78	82	85	TIDAK ADA	NEGERI
106	80	88	78	82	96	95	TIDAK ADA	NEGERI
107	82	86	81	81	85	77	TIDAK ADA	NEGERI
108	79	86	80	81	85	81	TIDAK ADA	NEGERI
109	82	83	78	87	95	81	TIDAK ADA	NEGERI
110	85	81	82	87	88	79	TIDAK ADA	NEGERI
111	84	80	82	82	81	83	TIDAK ADA	NEGERI
112	82	84	78	80	90	83	TIDAK ADA	NEGERI
113	80	82	78	85	91	85	TIDAK ADA	NEGERI
114	88	93	88	93	98	90	ADA	NEGERI
115	85	82	81	85	79	80	TIDAK ADA	NEGERI
116	83	87	86	81	91	93	TIDAK ADA	NEGERI
117	84	80	81	88	82	79	TIDAK ADA	NEGERI
118	88	93	83	87	81	78	TIDAK ADA	NEGERI
119	81	83	82	82	82	81	TIDAK ADA	NEGERI
120	81	83	83	81	84	85	TIDAK ADA	NEGERI
121	91	90	92	91	90	92	TIDAK ADA	NEGERI
122	97	87	82	77	81	81	TIDAK ADA	NEGERI
123	87	85	93	91	87	91	TIDAK ADA	NEGERI
124	83	80	81	83	81	82	TIDAK ADA	NEGERI
125	85	92	91	91	89	91	TIDAK ADA	NEGERI
126	95	91	94	97	91	98	TIDAK ADA	NEGERI
127	90	90	83	89	92	84	TIDAK ADA	NEGERI
128	95	78	84	92	89	84	TIDAK ADA	NEGERI

129	83	81	80	82	80	83	TIDAK ADA	NEGERI
130	84	88	85	90	91	92	TIDAK ADA	NEGERI
131	81	85	77	85	87	85	TIDAK ADA	NEGERI
132	87	87	82	90	88	77	TIDAK ADA	NEGERI
133	81	80	82	83	83	91	TIDAK ADA	NEGERI
134	87	80	83	85	84	88	TIDAK ADA	NEGERI

Tabel 3.2 Data uji

NO	NILAI RAPOR			NILAI TRY OUT			AKADMIK	SEKOLAH
	IPA	BINDO	MTK	IPA	BINDO	MTK		
61	83	82	70	85	85	79	TIDAK ADA	SWASTA
62	85	75	84	78	86	79	TIDAK ADA	SWASTA
63	75	78	71	81	80	79	TIDAK ADA	SWASTA
64	76	78	71	78	78	71	TIDAK ADA	SWASTA
65	73	78	71	76	78	70	TIDAK ADA	SWASTA
66	84	83	80	83	83	82	TIDAK ADA	SWASTA
67	73	78	70	73	78	70	TIDAK ADA	SWASTA
68	82	83	80	84	83	80	TIDAK ADA	SWASTA
69	80	78	72	87	81	71	TIDAK ADA	SWASTA
70	80	78	71	73	83	71	TIDAK ADA	SWASTA
71	81	81	70	81	78	71	TIDAK ADA	SWASTA
72	79	83	85	83	83	80	TIDAK ADA	SWASTA
73	75	83	70	83	83	70	TIDAK ADA	SWASTA
74	73	78	71	85	78	72	TIDAK ADA	SWASTA
135	90	94	91	94	91	95	ADA	NEGERI
136	80	84	84	79	89	83	TIDAK ADA	NEGERI
137	81	85	83	86	88	86	TIDAK ADA	NEGERI
138	79	81	82	87	85	85	TIDAK ADA	NEGERI
139	83	91	85	90	91	90	TIDAK ADA	NEGERI
140	84	85	88	89	88	96	TIDAK ADA	NEGERI
141	90	84	90	88	88	98	TIDAK ADA	NEGERI
142	84	81	84	84	84	91	TIDAK ADA	NEGERI

Tabel 3.3 Data Atribut

No	Atribut	Keterangan	Tipe
1	IPA 1	Nilai Ujian Nasional	Numerik
2	BINDO 1	Nilai Ujian Nasional	Numerik

3	MTK 1	Nilai Ujian Nasional	Numerik
4	IPA 2	Nilai <i>Try out</i>	Numerik
5	BIN 2	Nilai <i>Try out</i>	Numerik
6	MTK 2	Nilai <i>Try out</i>	Numerik
7	AKADEMIK	Prestasi Olimpiade akademik	Kategorikal

### 3.4 Perhitungan *Decision Tree C4.5*

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) Ilmu Pengetahuan Alam (IPA1)  
Posisi  $v$  yang digunakan pada atribut IPA adalah nilai antara (70, 80, 90, 100)
- c) Bahasa Indonesia 1 (BIN1)  
Posisi  $v$  yang digunakan pada atribut BIN1 adalah nilai antara (70, 80, 90, 100).
- d) Matematika 1 (MTK1)  
Posisi  $v$  yang digunakan pada atribut MTK1 adalah nilai antara (70,80,90,100).
- e) Ilmu Pengetahun Alam 2 (IPA2)  
Posisi  $v$  yang digunakan pada atribut IPA2 adalah nilai antara (70, 80, 90,100).
- f) Bahasa Indonesia 2 (BIN2)  
Posisi  $v$  yang digunakan pada atribut BIN2 adalah nilai antara (70, 80, 90, 100).

## g) Matematika 2 (MTK2)

Posisi  $v$  yang digunakan pada atribut BIN2 adalah nilai antara (70, 80 , 90,100).

## h) AKADEMIK (AKDM)

Akademik adalah atribut yang bernilai kategorikal sehingga tidak perlu ditentukan nilai posisi  $v$ .

## i) Kelas

Pada variabel diagnosa yang kelas klasifikasi yang digunakan adalah Negeri dan Swasta.

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.

$$\begin{aligned} Entropy(S) &= -\frac{60}{120} * \log_2\left(\frac{60}{120}\right) - \frac{60}{120} * \log_2\left(\frac{60}{120}\right) \\ &= 0,5 + 0,5 = 1 \end{aligned}$$

Selanjutnya menghitung nilai gain untuk setiap atribut, jika atribut memiliki nilai numerik maka akan dilakukan perhitungan untuk menentukan posisi  $v$ . Berikut adalah perhitungan untuk menentukan posisi gain tertinggi untuk masing-masing posisi.

**Tabel 3.4** Posisi  $v$  untuk pemecahan atribu IPA 1

IPA 1	70				80			
	<=		>		<=		>	
NEGERI	0	0	60	0.5000	5	0.3080	55	0.1710
SWASTA	0	0	60	0.5000	52	0.1208	8	0.3781
sum	0		120		57		63	
entropi	0		1		0.4288		0.5491	
gain	0				0.5080			

**Tabel 3.4** Posisi  $v$  untuk pemecahan atribut IPA1 (lanjutan)

IPA 1	90				100			
	$\leq$		$>$		$\leq$		$>$	
NEGERI	51	1	9	0.0000	60	0.5000	0	0.0000
SWASTA	60	0	0	0.0000	60	0.5000	0	0.0000
sum	111		9		120		0	
entropi	0.9953		0		1		0	
gain	0.0793				0.0000			

Hasil penentuan posisi  $v$  pada atribut ipa1 menunjukkan nilai *gain* tertinggi didapat pada posisi  $v = 80$ . Maka untuk atribut ipa1 dilakukan diskretisasi pada  $v = 80$  ketika menghitung *entropy* dan *gain* pada semua atribut.

**Tabel 3.5** Posisi  $v$  untuk pemecahan atribu BIN 1

BINDO 1	70				80			
	$\leq$		$>$		$\leq$		$>$	
NEGERI	0	0	60	0.5	9	0.414	51	0.216
SWASTA	0	0	60	0.5	50	0.202	10	0.428
sum	0		120		59		61	
entropi	0		1		0.6162		0.6436	
gain	0				0.3699			

**Tabel 3.5** Posisi  $v$  untuk pemecahan atribu BIN 1 (lanjutan)

BINDO 1	90				100			
	$\leq$		$>$		$\leq$		$>$	
NEGERI	53	1	7	0	60	0.5	0	0
SWASTA	60	0	0	0	60	0.5	0	0
sum	113		7		120		0	
entropi	0.9972		0		1		0	
gain	0.061				0			

Hasil penentuan posisi  $v$  pada atribut bin1 menunjukkan nilai *gain* tertinggi didapat pada posisi  $v=80$ . Maka atribut bin1 dilakukan diskretisasi pada  $v=80$  ketika menghitung *entropy* dan *gain* pada semua atribut.

**Tabel 3.6** Posisi v untuk pemecahan atribu MTK 1

MTK 1	70				80			
	<=		>		<=		>	
NEGERI	0	0	60	0.355	12	0.439	48	0.082
SWASTA	35	0	25	0.519	57	0.228	3	0.24
sum	35		85		69		51	
entropi	0		0.874		0.6666		0.3228	
gain	0.3809				0.4795			

**Tabel 3.6** Posisi v untuk pemecahan atribu MTK 1 (lanjutan)

MTK 1	90				100			
	<=		>		<=		>	
NEGERI	55	1	5	0.347	60	0.5	0	0
SWASTA	58	0	2	0.516	60	0.5	0	0
sum	113		7		120		0	
entropi	0.9995		0.8631		1		0	
gain	0.0085				0			

Hasil penentuan posisi v pada atribut mtk1 menunjukkan nilai *gain* tertinggi didapat pada posisi  $v=80$ . Maka atribut mtk1 dilakukan diskretisasi pada  $v=80$  ketika menghitung *entropy* dan *gain* pada semua atribut.

**Tabel 3.7** Posisi v untuk pemecahan atribut IPA2

IPA 2	70				80			
	<=		>		<=		>	
NEGERI	0	0	60	0.5	4	0.272	56	0.169
SWASTA	0	0	60	0.5	52	0.099	8	0.375
sum	0		120		56		64	
entropi	0		1		0.3712		0.5436	
gain	0				0.5369			

**Tabel 3.7** Posisi v untuk pemecahan atribut IPA2 (lanjutan)

IPA 2	90				100			
	<=		>		<=		>	
NEGERI	53	1	7	0	60	0.5	0	0
SWASTA	60	0	0	0	60	0.5	0	0
sum	113		7		120		0	
entropi	0.9972		0		1		0	
gain	0.061				0			

Hasil penentuan posisi  $v$  pada atribut ipa2 menunjukkan nilai *gain* tertinggi didapat pada posisi  $v = 80$ . Maka untuk atribut ipa2 dilakukan diskretisasi pada  $v = 80$  ketika menghitung *entropy* dan *gain* pada semua atribut.

**Tabel 3.8** Posisi  $v$  untuk pemecahan atribut BIN2

BINDO 2	70				80			
	<=		>		<=		>	
NEGERI	0	0	60	0.5	4	0.319	56	0.334
SWASTA	0	0	60	0.5	39	0.128	21	0.511
sum	0		120		43		77	
entropi	0		1		0.4465		0.8454	
gain	0				0.2975			

**Tabel 3.8** Posisi  $v$  untuk pemecahan atribut BIN2 (lanjutan)

BINDO 2	90				100			
	<=		>		<=		>	
NEGERI	50	1	10	0.219	60	0.5	0	0
SWASTA	58	0	2	0.431	60	0.5	0	0
sum	108		12		120		0	
entropi	0.996		0.65		1		0	
gain	0.0386				0			

Hasil penentuan posisi  $v$  pada atribut bin2 menunjukkan nilai *gain* tertinggi didapat pada posisi  $v = 80$ . Maka untuk atribut bin2 dilakukan diskretisasi pada  $v = 80$  ketika menghitung *entropy* dan *gain* pada semua atribut.

**Tabel 3.9** Posisi  $v$  untuk pemecahan atribut MTK2

MTK 2	70				80			
	<=		>		<=		>	
NEGERI	0	0	60	0.429	6	0.326	54	0.096
SWASTA	23	0	37	0.53	56	0.133	4	0.266
sum	23		97		62		58	
entropi	0		0.9591		0.4587		0.3621	
gain	0.2247				0.5880			

**Tabel 3.9** Posisi v untuk pemecahan atribut MTK2 (lanjutan)

MTK 2	90				100			
	<=		>		<=		>	
NEGERI	46	1	14	0	60	0.5	0	0
SWASTA	60	0	0	0	60	0.5	0	0
sum	106		14		120		0	
entropi	0.9874		0		1		0	
gain	0.1278				0.0000			

Hasil penentuan posisi v pada atribut mtk2 menunjukkan nilai *gain* tertinggi didapat pada posisi  $v = 80$ . Maka untuk atribut mtk2 dilakukan diskretisasi pada  $v = 80$  ketika menghitung *entropy* dan *gain* pada semua atribut.

Selanjutnya menghitung *entropy* dan *gain* untuk atribut prestasi akademik, perhitungan untuk atribut ini tidak perlu menggunakan posisi v karena atribut memiliki nilai kategorikal. Berikut adalah perhitungannya.

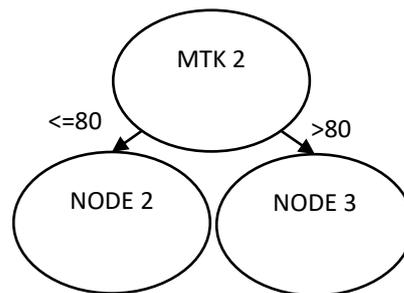
$$\begin{aligned}
 \text{Gain}(\text{akademik}) &= 1 - \left( \frac{2}{120} * \left( -\frac{2}{2} * \log_2 \left( \frac{2}{2} \right) - \frac{0}{2} * \log_2 \left( \frac{0}{2} \right) \right) \right. \\
 &\quad \left. + \frac{118}{120} * \left( -\frac{58}{118} * \log_2 \left( \frac{58}{118} \right) - \frac{60}{118} * \log_2 \left( \frac{60}{118} \right) \right) \right) = 0,0169
 \end{aligned}$$

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 table 3.10 **Tabel 3.10** Hasil perhitungfan entropy dan gain untuk node akar

Perhitungan entropy dan gain		NODE 1				
		Jml	Negeri	Swasta	Entropy	Gain
<b>Total</b>		120	60	60	1	
<b>IPA1</b>	<=80	57	5	52	0,4288	0,5080
	>80	63	55	8	0,5491	
<b>BIN1</b>	<=80	59	9	50	0,6162	03699
	>80	61	51	10	0,6436	
<b>MTK1</b>	<=80	69	12	57	0,6666	0,4795
	>80	51	48	3	0,3228	

<b>IPA2</b>	<b>&lt;=80</b>	56	4	52	0,3712	0,5369
	<b>&gt;80</b>	64	56	8	0,5436	
<b>BIN2</b>	<b>&lt;=80</b>	43	4	39	0,4465	0,2975
	<b>&gt;80</b>	77	56	21	0,8454	
<b>MTK2</b>	<b>&lt;=80</b>	62	6	54	0,4587	0,5880
	<b>&gt;80</b>	58	54	4	0,3621	
<b>AKDM</b>	<b>ADA</b>	2	2	0	0	0,0169
	<b>TIDAK ADA</b>	118	58	68	0,9998	

Hasil yang didapat di tabel 3.10 menunjukkan bahwa *gain* tertinggi ada di atribut MTK2 , maka MTK2 dijadikan sebagai *node* akar. Pada atribut MTK2 *entropy* yang dihasilkan pada nilai  $\leq 80$  sehingga atribut MTK2 dijadikan syarat kondisi di node 2 dan node 3. Pohon yang terbentuk untuk proses pada node 1 adalah sebagai berikut.



**Gambar 3.3** Pohon Keputusan yang Terbentuk pada Node 1

**Tabel 3.11** Data Kasus pada MTK2  $\leq 80$

NO	NILAI RAPOR			NILAI TRY OUT			AKADMIK	SEKOLAH
	IPA	BINDO	MTK	IPA	BINDO	MTK		
1	85	78	81	82	78	80	TIDAK ADA	SWASTA
2	73	78	70	74	78	71	TIDAK ADA	SWASTA
3	73	78	70	73	78	70	TIDAK ADA	SWASTA
4	74	78	70	73	78	70	TIDAK ADA	SWASTA
5	78	78	70	74	78	70	TIDAK ADA	SWASTA
6	73	78	70	73	78	71	TIDAK ADA	SWASTA
7	73	78	70	73	78	70	TIDAK ADA	SWASTA
8	73	78	70	74	78	71	TIDAK ADA	SWASTA
9	80	78	70	73	78	71	TIDAK ADA	SWASTA
10	73	78	70	74	78	70	TIDAK ADA	SWASTA
11	86	78	74	76	79	72	TIDAK ADA	SWASTA
12	79	78	71	84	81	78	TIDAK ADA	SWASTA
13	84	78	70	74	79	71	TIDAK ADA	SWASTA

14	73	78	70	73	78	70	TIDAK ADA	SWASTA
15	73	78	70	75	78	71	TIDAK ADA	SWASTA
16	75	78	77	74	78	77	TIDAK ADA	SWASTA
17	74	79	70	73	78	73	TIDAK ADA	SWASTA
18	79	79	76	78	79	76	TIDAK ADA	SWASTA
19	73	78	70	74	78	70	TIDAK ADA	SWASTA
20	73	78	70	73	78	70	TIDAK ADA	SWASTA
21	73	78	70	73	78	70	TIDAK ADA	SWASTA
22	76	78	71	75	83	78	TIDAK ADA	SWASTA
23	78	78	72	78	78	74	TIDAK ADA	SWASTA
24	73	78	70	75	80	73	TIDAK ADA	SWASTA
25	75	78	71	74	79	71	TIDAK ADA	SWASTA
26	83	81	75	78	88	75	TIDAK ADA	SWASTA
27	74	78	70	74	90	72	TIDAK ADA	SWASTA
28	76	83	70	89	95	80	TIDAK ADA	SWASTA
29	76	81	76	80	86	80	TIDAK ADA	SWASTA
30	78	78	70	89	92	72	TIDAK ADA	SWASTA
31	78	78	72	77	87	73	TIDAK ADA	SWASTA
32	76	79	77	83	84	80	TIDAK ADA	SWASTA
34	76	78	70	74	78	70	TIDAK ADA	SWASTA
35	73	78	91	74	78	75	TIDAK ADA	SWASTA
36	76	78	70	74	78	70	TIDAK ADA	SWASTA
37	83	85	70	80	87	73	TIDAK ADA	SWASTA
41	74	78	70	75	78	70	TIDAK ADA	SWASTA
42	76	85	70	75	81	70	TIDAK ADA	SWASTA
43	83	85	78	87	88	78	TIDAK ADA	SWASTA
44	74	78	72	75	79	72	TIDAK ADA	SWASTA
45	73	78	70	74	78	70	TIDAK ADA	SWASTA
46	76	80	71	76	79	70	TIDAK ADA	SWASTA
47	74	82	70	78	84	76	TIDAK ADA	SWASTA
48	73	78	70	76	79	70	TIDAK ADA	SWASTA
49	74	78	70	73	78	70	TIDAK ADA	SWASTA
50	73	80	70	73	78	70	TIDAK ADA	SWASTA
51	74	78	80	74	81	73	TIDAK ADA	SWASTA
52	73	78	70	73	78	70	TIDAK ADA	SWASTA
53	85	78	70	80	84	70	TIDAK ADA	SWASTA
54	73	78	70	73	80	70	TIDAK ADA	SWASTA
55	77	80	73	77	81	73	TIDAK ADA	SWASTA
56	73	83	74	83	87	75	TIDAK ADA	SWASTA
57	73	78	70	73	78	70	TIDAK ADA	SWASTA
58	73	78	71	75	80	75	TIDAK ADA	SWASTA

59	73	78	71	73	78	70	TIDAK ADA	SWASTA
60	81	86	72	79	84	77	TIDAK ADA	SWASTA
107	82	86	81	81	85	77	TIDAK ADA	NEGERI
110	85	81	82	87	88	79	TIDAK ADA	NEGERI
115	85	82	81	85	79	80	TIDAK ADA	NEGERI
117	84	80	81	88	82	79	TIDAK ADA	NEGERI
118	88	93	83	87	81	78	TIDAK ADA	NEGERI
132	87	87	82	90	88	77	TIDAK ADA	NEGERI

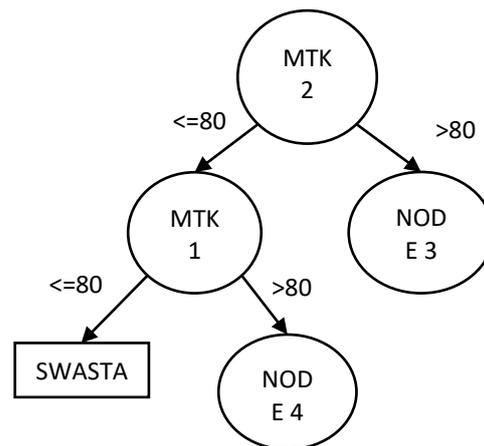
Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk atribut MTK2  $\leq 80$  saja. Hasil pemecahan tersebut menyisakan 62 data yang terdiri dari 56 data dengan kelas prediksi Swasta dan 6 data dengan kelas prediksi Negeri. Hasil 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 table 3.12

**Tabel 3.12** Hasil perhitungan entropy dan gain untuk node 2.

Perhitungan entropy dan gain		NODE 2				
Total		Jml	Negeri	Swasta	Entropy	Gain
IPA1	$\leq 80$	48	0	48	0	0,2362
	$> 80$	14	6	8	0,9852	
BIN1	$\leq 80$	48	1	47	0,1461	0,1333
	$> 80$	14	5	9	0,9403	
MTK1	$\leq 80$	54	0	54	0	0,3540
	$> 80$	8	6	2	0,8113	
IPA2	$\leq 80$	49	0	49	0	0,2499
	$> 80$	13	6	7	0,9957	
BIN2	$\leq 80$	39	1	38	0,1720	0,0703
	$> 80$	23	5	18	0,7554	
MTK2	$\leq 70$	23	0	23	0	0,0691
	$> 70$	39	6	33	0,6194	
AKDM	ADA	0	0	0	0	0
	TIDAK ADA	62	6	56	0,4587	

Hasil yang didapat di tabel 3.12 menunjukkan bahwa *gain* tertinggi ada di atribut MTK1, maka MTK1 dijadikan sebagai *node* internal

(node 2). Pada atribut MTK1 *entropy* yang dihasilkan pada nilai  $\leq 80$  memiliki nilai 0 sehingga dapat dipastikan semua data latih yang memiliki MTK  $1 \leq 80$  masuk dalam kelas prediksi Swasta. Sedangkan nilai *entropy* yang  $> 80$  akan dijadikan sebagai node selanjutnya. Pohon yang terbentuk untuk proses pada node 2 adalah sebagai berikut.



**Gambar 3.4** Pohon Keputusan yang Terbentuk pada Node 2

**Tabel 3.13** Data Kasus pada MTK2  $> 80$

NO	NILAI RAPOR			NILAI TRY OUT			AKADMIK	SEKOLAH
	IPA	BINDO	MTK	IPA	BINDO	MTK		
33	80	78	92	75	78	85	TIDAK ADA	SWASTA
38	74	78	70	74	81	82	TIDAK ADA	SWASTA
39	75	78	73	74	85	82	TIDAK ADA	SWASTA
40	76	83	73	85	84	82	TIDAK ADA	SWASTA
75	91	84	82	84	84	85	TIDAK ADA	NEGERI
76	83	84	84	87	86	85	TIDAK ADA	NEGERI
77	91	78	80	87	88	95	TIDAK ADA	NEGERI
78	94	78	89	83	85	91	TIDAK ADA	NEGERI
79	94	83	83	81	84	85	TIDAK ADA	NEGERI
80	90	81	76	89	87	91	TIDAK ADA	NEGERI
81	92	84	85	88	89	93	TIDAK ADA	NEGERI
82	86	83	83	84	84	83	TIDAK ADA	NEGERI
83	90	82	81	80	82	85	TIDAK ADA	NEGERI
84	83	83	81	87	84	84	TIDAK ADA	NEGERI
85	87	83	85	82	86	85	TIDAK ADA	NEGERI
86	89	81	91	90	86	97	TIDAK ADA	NEGERI
87	82	83	82	82	85	85	TIDAK ADA	NEGERI
88	85	95	83	83	78	83	TIDAK ADA	NEGERI

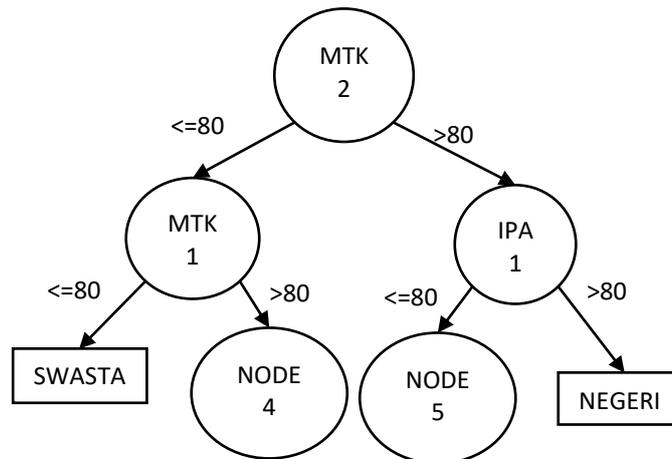
89	83	84	83	83	82	81	TIDAK ADA	NEGERI
90	81	83	82	82	86	85	TIDAK ADA	NEGERI
91	87	85	82	82	86	85	TIDAK ADA	NEGERI
92	83	84	82	83	84	83	TIDAK ADA	NEGERI
93	85	83	80	83	83	83	TIDAK ADA	NEGERI
94	86	90	83	87	90	90	TIDAK ADA	NEGERI
95	84	84	83	85	87	93	TIDAK ADA	NEGERI
96	84	81	83	85	84	83	TIDAK ADA	NEGERI
97	83	83	82	83	83	82	TIDAK ADA	NEGERI
98	85	86	83	84	84	83	TIDAK ADA	NEGERI
99	78	84	80	87	88	83	TIDAK ADA	NEGERI
100	84	92	82	84	92	82	TIDAK ADA	NEGERI
101	89	93	78	93	92	86	ADA	NEGERI
102	83	83	82	84	83	82	TIDAK ADA	NEGERI
103	83	78	84	84	78	84	TIDAK ADA	NEGERI
104	81	84	83	81	83	83	TIDAK ADA	NEGERI
105	80	82	83	78	82	85	TIDAK ADA	NEGERI
106	80	88	78	82	96	95	TIDAK ADA	NEGERI
108	79	86	80	81	85	81	TIDAK ADA	NEGERI
109	82	83	78	87	95	81	TIDAK ADA	NEGERI
111	84	80	82	82	81	83	TIDAK ADA	NEGERI
112	82	84	78	80	90	83	TIDAK ADA	NEGERI
113	80	82	78	85	91	85	TIDAK ADA	NEGERI
114	88	93	88	93	98	90	ADA	NEGERI
116	83	87	86	81	91	93	TIDAK ADA	NEGERI
119	81	83	82	82	82	81	TIDAK ADA	NEGERI
120	81	83	83	81	84	85	TIDAK ADA	NEGERI
121	91	90	92	91	90	92	TIDAK ADA	NEGERI
122	97	87	82	77	81	81	TIDAK ADA	NEGERI
123	87	85	93	91	87	91	TIDAK ADA	NEGERI
124	83	80	81	83	81	82	TIDAK ADA	NEGERI
125	85	92	91	91	89	91	TIDAK ADA	NEGERI
126	95	91	94	97	91	98	TIDAK ADA	NEGERI
127	90	90	83	89	92	84	TIDAK ADA	NEGERI
128	95	78	84	92	89	84	TIDAK ADA	NEGERI
129	83	81	80	82	80	83	TIDAK ADA	NEGERI
130	84	88	85	90	91	92	TIDAK ADA	NEGERI
131	81	85	77	85	87	85	TIDAK ADA	NEGERI
133	81	80	82	83	83	91	TIDAK ADA	NEGERI
134	87	80	83	85	84	88	TIDAK ADA	NEGERI

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk atribut MTK3 >70 saja. Hasil pemecahan tersebut menyisakan 58 data yang terdiri dari 4 data dengan kelas prediksi Swasta dan 54 data dengan kelas prediksi Negeri. Hasil perhitungan semua atribut diperoleh entropy dan gain tertinggi maka dihitung gain untuk masing-masing atribut dipilih dengan nilai yang paling besar pula. Hasilnya disajikan pada table 3.14

**Tabel 3.14** Hasil perhitungfan entropy dan gain untuk node 3.

Perhitungan entropy dan gain		NODE 3				
		Jml	Negeri	Swasta	Entropy	Gain
<b>Total</b>		58	4	54	0,3620	
<b>IPA1</b>	<b>&lt;=80</b>	9	5	4	0,9911	0,2083
	<b>&gt;80</b>	49	49	0	0	
<b>BIN1</b>	<b>&lt;=80</b>	11	8	3	0,8454	0,0814
	<b>&gt;80</b>	47	46	1	0,1484	
<b>MTK1</b>	<b>&lt;=80</b>	15	12	3	0,7219	0,0572
	<b>&gt;80</b>	43	42	1	0,1594	
<b>IPA2</b>	<b>&lt;=80</b>	7	4	3	0,9852	0,1207
	<b>&gt;80</b>	51	50	1	0,1392	
<b>BIN2</b>	<b>&lt;=90</b>	48	44	4	0,4138	0,0196
	<b>&gt;90</b>	10	10	0	0	
<b>MTK2</b>	<b>&lt;=90</b>	14	14	0	0	0,0286
	<b>&gt;90</b>	44	40	4	0,4395	
<b>AKDM</b>	<b>ADA</b>	2	2	0	0	0,0036
	<b>TIDAK ADA</b>	56	52	4	0,3712	

Hasil yang didapat di tabel 3.14 menunjukkan bahwa *gain* tertinggi ada di atribut IPA1 maka IPA1 dijadikan sebagai *node* internal (node 3). Pada atribut IPA1 *entropy* yang dihasilkan pada nilai >80 memiliki nilai 0 sehingga dapat dipastikan semua data latih yang beratribut IPA1 >80 masuk dalam kelas prediksi Negeri. Sedangkan nilai *entropy* yang <80 akan dijadikan sebagai *node* selanjutnya. Pohon yang terbentuk untuk proses pada node 3 adalah sebagai berikut.



**Gambar 3.5** Pohon Keputusan yang Terbentuk pada Node 3

**Tabel 3.15** Data Kasus pada MTK1 >80

NO	NILAI RAPOR			NILAI TRY OUT			AKADMIK	SEKOLAH
	IPA	BINDO	MTK	IPA	BINDO	MTK		
1	85	78	81	82	78	80	TIDAK ADA	SWASTA
35	73	78	91	74	78	75	TIDAK ADA	SWASTA
107	82	86	81	81	85	77	TIDAK ADA	NEGERI
110	85	81	82	87	88	79	TIDAK ADA	NEGERI
115	85	82	81	85	79	80	TIDAK ADA	NEGERI
117	84	80	81	88	82	79	TIDAK ADA	NEGERI
118	88	93	83	87	81	78	TIDAK ADA	NEGERI
132	87	87	82	90	88	77	TIDAK ADA	NEGERI

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk atribut MTK1 >80 saja. Hasil pemecahan tersebut menyisakan 8 data yang terdiri dari 2 data dengan kelas prediksi Swasta dan 6 data dengan kelas prediksi Negeri. Hasil perhitungan semua atribut diperoleh entropy dan gain tertinggi maka dihitung gain untuk masing-masing atribut dipilih dengan nilai yang paling besar pula. Hasilnya disajikan pada table 3.16

**Tabel 3.16** Hasil perhitungfan entropy dan gain untuk node 4.

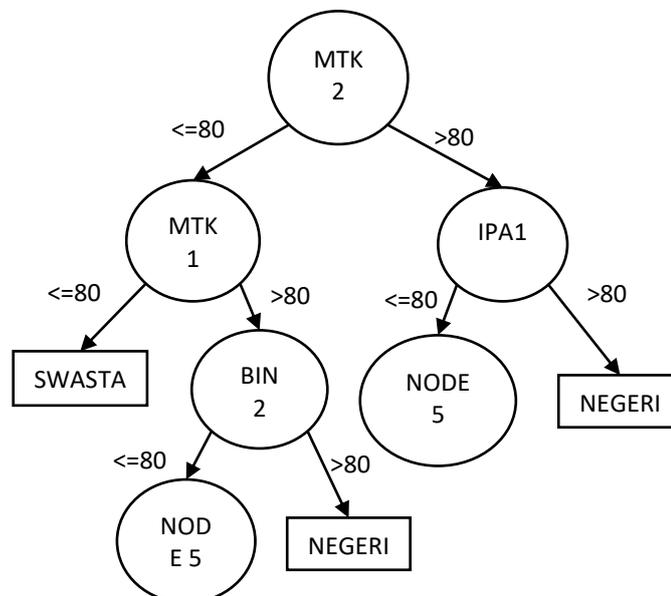
**Perhitungan entropy dan gain**

**NODE 4**

	Jml	Negeri	Swasta	Entropy	Gain
<b>Total</b>	8	6	2	0,8113	

<b>IPA1</b>	<b>&lt;=80</b>	1	0	1	0	0,2935
	<b>&gt;80</b>	7	6	1	0,5917	
<b>BIN1</b>	<b>&lt;=80</b>	3	1	3	0,9183	0,4669
	<b>&gt;80</b>	5	5	0	0	
<b>MTK1</b>	<b>&lt;=90</b>	7	6	1	0,5917	0,2935
	<b>&gt;90</b>	1	0	1	0	
<b>IPA2</b>	<b>&lt;=80</b>	1	0	1	0	0,2935
	<b>&gt;80</b>	7	6	1	0,5917	
<b>BIN2</b>	<b>&lt;=80</b>	3	1	2	0,9183	0,4669
	<b>&gt;80</b>	5	5	0	0	

Hasil yang didapat di tabel 3.16 menunjukkan bahwa *gain* tertinggi ada di dua atribut yaitu atribut BIN2 dan BIN1, boleh pilih salah satunya maka BIN2 dijadikan sebagai *node* internal (node 4). Pada atribut BIN2 *entropy* yang dihasilkan pada nilai >80 memiliki nilai 0 sehingga dapat dipastikan semua data latih yang memiliki BIN >80 masuk dalam kelas prediksi Negeri. Sedangkan nilai *entropy* yang <=80 masuk perhitungan node selanjutnya. Pohon yang terbentuk untuk proses pada node 4 adalah sebagai berikut.



**Gambar 3.6** Pohon Keputusan yang Terbentuk pada Node 4

**Tabel 3.17** Data Kasus pada IPA1  $\leq 80$ 

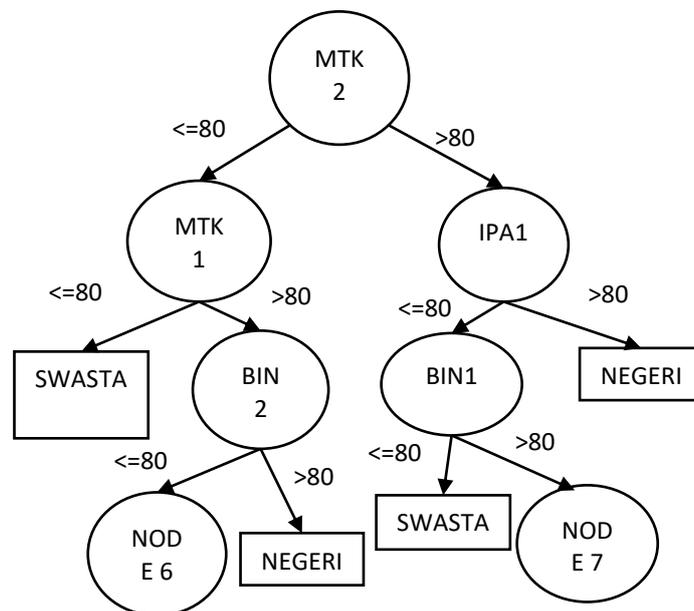
NO	NILAI RAPOR			NILAI TRY OUT			AKADMIK	SEKOLAH
	IPA	BINDO	MTK	IPA	BINDO	MTK		
33	80	78	92	75	78	85	TIDAK ADA	SWASTA
38	74	78	70	74	81	82	TIDAK ADA	SWASTA
39	75	78	73	74	85	82	TIDAK ADA	SWASTA
40	76	83	73	85	84	82	TIDAK ADA	SWASTA
99	78	84	80	87	88	83	TIDAK ADA	NEGERI
105	80	82	83	78	82	85	TIDAK ADA	NEGERI
106	80	88	78	82	96	95	TIDAK ADA	NEGERI
108	79	86	80	81	85	81	TIDAK ADA	NEGERI
113	80	82	78	85	91	85	TIDAK ADA	NEGERI

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk atribut IPA1  $\leq 80$  saja. Hasil pemecahan tersebut menyisakan 9 data yang terdiri dari 4 data dengan kelas prediksi Swasta dan 5 data dengan kelas prediksi Negeri. Hasil perhitungan semua atribut diperoleh entropy dan gain tertinggi maka dihitung gain untuk masing-masing atribut dipilih dengan nilai yang paling besar pula. Hasilnya disajikan pada table 3.18

**Tabel 3.18** Hasil perhitungan entropy dan gain untuk node 5.

Perhitungan entropy dan gain			NODE 5			
		Jml	Negeri	Swasta	Entropy	Gain
<b>Total</b>		9	5	4	0,9911	
<b>BIN1</b>	$\leq 80$	3	0	3	0	0,5577
	$> 80$	6	5	1	0,65	
<b>MTK1</b>	$\leq 90$	8	5	3	0,9544	0,1427
	$> 90$	1	0	1	0	
<b>IPA2</b>	$\leq 80$	4	1	3	0,8113	0,2294
	$> 80$	5	4	1	0,7219	
<b>BIN2</b>	$\leq 90$	7	3	4	0,9852	0,2248
	$> 90$	2	2	0	0	
<b>MTK2</b>	$\leq 90$	8	4	4	1	0,1022
	$> 90$	1	1	0	0	

Hasil yang didapat di tabel 3.18 menunjukkan bahwa *gain* tertinggi ada di atribut BIN1, maka BIN1 dijadikan sebagai *node* internal (node 5). Pada atribut BIN1 *entropy* yang dihasilkan pada nilai  $\leq 80$  yang memiliki nilai 0 maka data latih yang beratribut  $BIN1 \leq 80$  masuk dalam kelas Swasta. Untuk data latih yang beratribut  $BIN1 > 80$  akan masuk dalam perhitungan *node* selanjutnya. Pohon yang terbentuk untuk *node* 5 adalah sebagai berikut.



**Gambar 3.7** Pohon Keputusan yang Terbentuk pada Node 5

**Tabel 3.19** Data Kasus pada BIN2  $\leq 80$

NO	NILAI RAPOR			NILAI TRY OUT			AKADMIK	SEKOLAH
	IPA	BINDO	MTK	IPA	BINDO	MTK		
1	85	78	81	82	78	80	TIDAK ADA	SWASTA
35	73	78	91	74	78	75	TIDAK ADA	SWASTA
115	85	82	81	85	79	80	TIDAK ADA	NEGERI

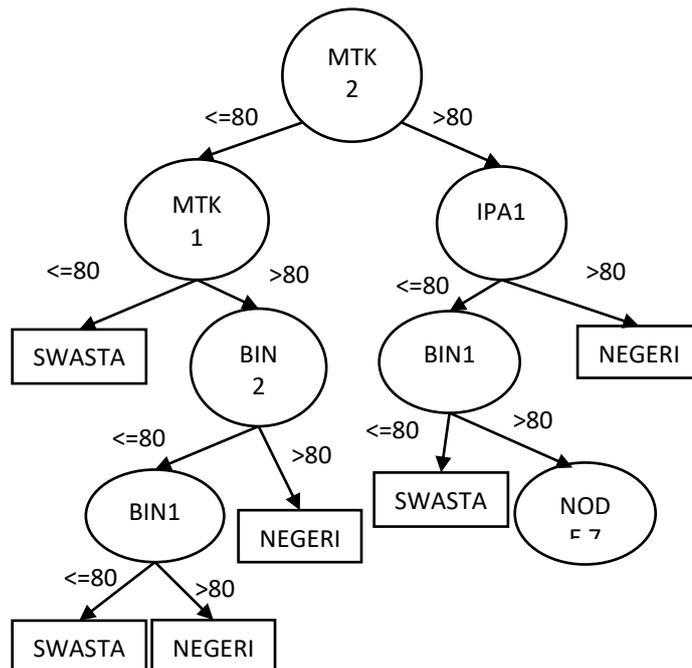
Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk atribut BIN2  $\leq 80$  saja. Hasil pemecahan tersebut menyisakan 3 data yang terdiri dari 2 data dengan kelas prediksi Swasta dan 1 data dengan kelas prediksi Negeri. Hasil perhitungan semua atribut diperoleh *entropy* dan *gain* tertinggi

maka dihitung gain untuk masing-masing atribut dipilih dengan nilai yang paling besar pula. Hasilnya disajikan pada table 3.23

**Tabel 3.20** Hasil perhitungan entropy dan gain untuk node 6.

Perhitungan entropy dan gain			NODE 6			
		Jml	Negeri	Swasta	Entropy	Gain
<b>Total</b>		3	1	2	0,9183	
<b>IPA1</b>	<b>&lt;=80</b>	1	0	1	0	0,2516
	<b>&gt;80</b>	2	1	1	1	
<b>BIN1</b>	<b>&lt;=80</b>	2	0	2	0	0,9183
	<b>&gt;80</b>	1	1	0	0	
<b>MTK1</b>	<b>&lt;=90</b>	2	1	1	1	0,2516
	<b>&gt;90</b>	1	0	1	0	
<b>IPA2</b>	<b>&lt;=80</b>	1	0	1	0	0,2516
	<b>&gt;80</b>	2	1	1	1	

Hasil yang didapat di tabel 3.20 menunjukkan bahwa *gain* tertinggi ada di atribut BIN1, maka BIN1 dijadikan sebagai *node* internal (node 6). Pada atribut BIN1 *entropy* yang dihasilkan pada nilai  $\leq 80$  dan  $> 80$  dengan nilai *entropy* 0 sehingga dapat dipastikan semua data latih yang memiliki  $BIN1 \leq 80$  masuk dalam kelas prediksi Swasta sedangkan data latih dengan atribut  $BIN1 > 80$  masuk dalam kelas Negeri. Pohon yang terbentuk untuk proses pada node 6 adalah sebagai berikut.



**Gambar 3.8** Pohon Keputusan yang Terbentuk pada Node 6**Tabel 3.21** Data Kasus pada BIN1>80

NO	NILAI RAPOR			NILAI TRY OUT			AKADEMIK	SEKOLAH
	IPA	BINDO	MTK	IPA	BINDO	MTK		
40	76	83	73	85	84	82	TIDAK ADA	SWASTA
99	78	84	80	87	88	83	TIDAK ADA	NEGERI
105	80	82	83	78	82	85	TIDAK ADA	NEGERI
106	80	88	78	82	96	95	TIDAK ADA	NEGERI
108	79	86	80	81	85	81	TIDAK ADA	NEGERI
113	80	82	78	85	91	85	TIDAK ADA	NEGERI

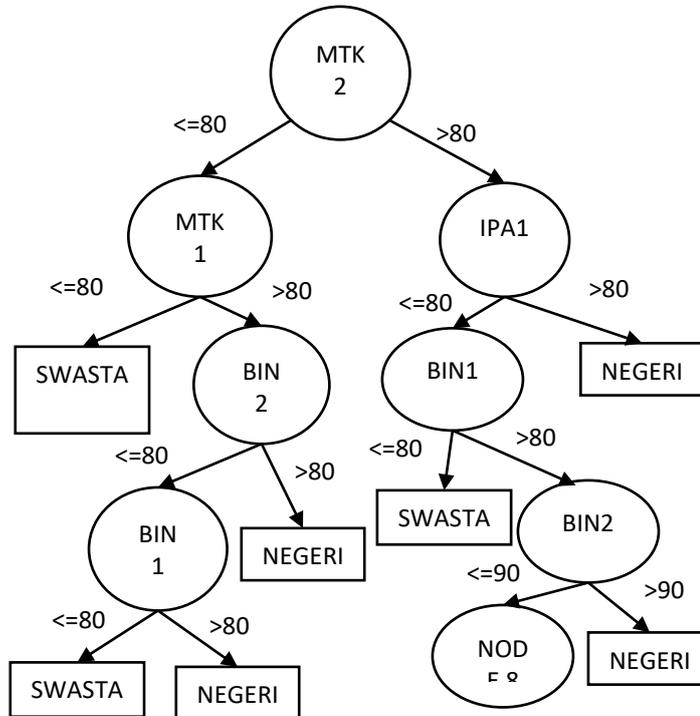
Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk atribut BIN1 >80 saja. Hasil pemecahan tersebut menyisakan 6 data yang terdiri dari 1 data dengan kelas prediksi Swasta dan 5 data dengan kelas prediksi Negeri. Hasil perhitungan semua atribut diperoleh entropy dan gain tertinggi maka dihitung gain untuk masing-masing atribut dipilih dengan nilai yang paling besar pula. Hasilnya disajikan pada table 3.22

**Tabel 3.22** Hasil perhitungannya entropy dan gain untuk node 7.

Perhitungan entropy dan gain				NODE 7		
Total		Jml	Negeri	Swasta	Entropy	Gain
MTK1	<=80	5	4	1	0,7219	0,0484
	>80	1	1	0	0	
IPA2	<=80	1	1	0	0	0,0484
	>80	5	4	1	0,7219	
BIN2	<=90	4	3	1	0,8113	0,1092
	>90	2	2	0	0	
MTK2	<=90	5	4	1	0,7219	0,0484
	>90	1	1	0	0	

Hasil yang didapat di tabel 3.22 menunjukkan bahwa *gain* tertinggi ada di atribut BIN2, maka BIN2 dijadikan sebagai *node* internal (node 7). Pada atribut BIN2 *entropy* yang dihasilkan pada nilai >90 yang memiliki nilai 0 sehingga dapat dipastikan semua data latih yang memiliki BIN2>90 masuk

dalam kelas prediksi Negeri sedangkan atribut  $BIN \leq 90$  akan dilanjutkan dalam perhitungan selanjutnya. Pohon yang terbentuk untuk proses pada node 5 adalah sebagai berikut.



. Gambar 3.9 Pohon Keputusan yang Terbentuk pada Node 7

Tabel 3.23 Data Kasus pada  $BIN2 \leq 90$

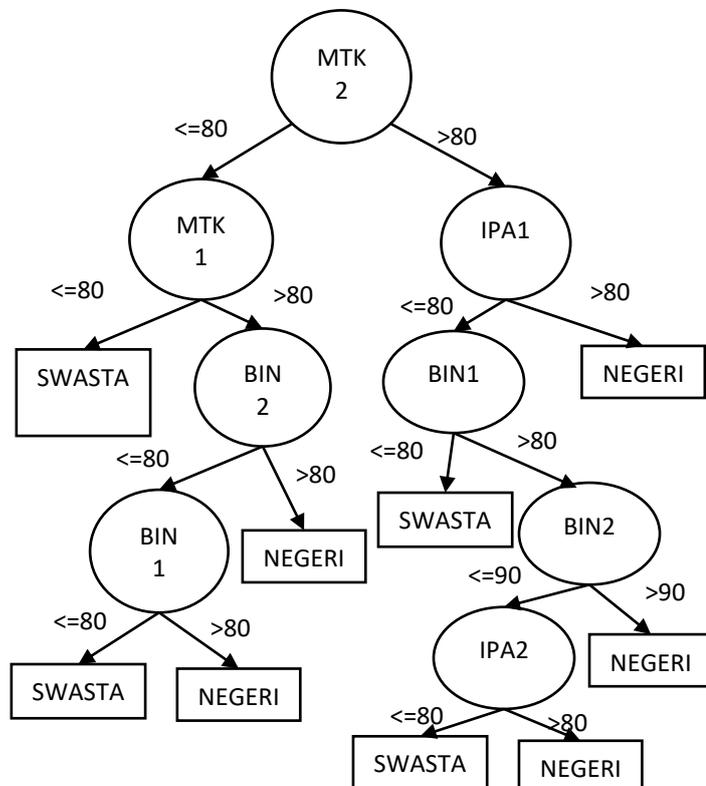
NO	NILAI RAPOR			NILAI TRY OUT			AKADEMIK	SEKOLAH
	IPA	BINDO	MTK	IPA	BINDO	MTK		
40	76	83	73	85	84	82	TIDAK ADA	SWASTA
99	78	84	80	87	88	83	TIDAK ADA	NEGERI
105	80	82	83	78	82	85	TIDAK ADA	NEGERI
108	79	86	80	81	85	81	TIDAK ADA	NEGERI

Perhitungan selanjutnya adalah dengan menggunakan data yang sudah mengalami pemecahan sehingga diperoleh data dengan nilai untuk atribut  $BIN2 \leq 90$  saja. Hasil pemecahan tersebut menyisakan 4 data yang terdiri dari 1 data dengan kelas prediksi Swasta dan 3 data dengan kelas prediksi Negeri. Hasil perhitungan semua atribut diperoleh entropy dan gain tertinggi maka dihitung gain untuk masing-masing atribut dipilih dengan nilai yang paling besar pula. Hasilnya disajikan pada table 3.2

**Tabel 3.24** Hasil perhitungfan entropy dan gain untuk node 8.

Perhitungan entropy dan gain		NODE 8				
		Jml	Negeri	Swasta	Entropy	Gain
<b>Total</b>		4	1	3	0,8113	
<b>MTK1</b>	$\leq 80$	3	2	1	0,9183	0,1226
	$> 80$	1	1	0	0	
<b>IPA2</b>	$\leq 80$	1	1	0	0	0,1226
	$> 80$	3	2	1	0,9183	

Hasil yang didapat di tabel 3.24 menunjukkan bahwa *gain* tertinggi ada di dua atribut yaitu MTK1 dan IPA2, maka IPA2 dijadikan sebagai *node* internal (node 8). Pada atribut IPA2 *entropy* yang dihasilkan pada nilai  $\leq 80$  memiliki nilai 0 sehingga dapat dipastikan semua data latih yang memiliki IPA2  $\leq 80$  masuk dalam kelas prediksi Swasta. Sedangkan nilai entrophy untuk atribut IPA2 $>80$  akan masuk dalam kelas prediksi Negeri karena sudah menjadi data heterogen. Pohon yang terbentuk untuk proses pada node 8 adalah sebagai berikut.



**Gambar 3.10** Pohon Keputusan yang Terbentuk pada Node 8

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

- RULE 1 : Jika nilai *try out* Matematika  $\leq 80$  dan nilai raport Matematika  $\leq 80$  maka diterima di sekolah Swasta.
- RULE 2 : Jika nilai *try out* Matematika  $\leq 80$  dan nilai raport Matematika  $> 80$  dan nilai *try out* Bahasa Indonesia  $> 80$  maka diterima di sekolah Negeri.
- RULE 3 : Jika nilai *try out* Matematika  $\leq 80$  dan nilai raport Matematika  $> 80$  dan nilai *try out* Bahasa Indonesia  $\leq 80$  dan nilai raport Bahasa Indonesia  $\leq 80$  maka diterima di sekolah Swasta.
- RULE 4 : Jika nilai *try out* Matematika  $\leq 80$  dan nilai raport Matematika  $> 80$  dan nilai *try out* Bahasa Indonesia  $\leq 80$  dan nilai raport Bahasa Indonesia  $> 80$  maka diterima di sekolah Negeri.
- RULE 5 : Jika nilai *try out* Matematika  $> 80$  dan nilai *try out* IPA  $> 80$  maka diterima di sekolah Negeri.
- RULE 6 : Jika nilai *try out* Matematika  $> 80$  dan nilai *try out* IPA  $\leq 80$  dan nilai *try out* Bahasa Indonesia  $\leq 80$  maka diterima di sekolah Swasta.
- RULE 7 : Jika nilai *try out* Matematika  $> 80$  dan nilai *try out* IPA  $\leq 80$  dan nilai *try out* Bahasa Indonesia  $> 80$  dan nilai raport Bahasa Indonesia  $> 90$  maka diterima di sekolah Negeri.
- RULE 8 : Jika nilai *try out* Matematika  $> 80$  dan nilai *try out* IPA  $\leq 80$  dan nilai *try out* Bahasa Indonesia  $> 80$  dan nilai raport Bahasa Indonesia  $\leq 90$  dan nilai raport IPA  $\leq 80$  maka diterima di sekolah Swasta.
- RULE 9 : Jika nilai *try out* Matematika  $> 80$  dan nilai *try out* IPA  $\leq 80$  dan nilai *try out* Bahasa Indonesia  $> 80$  dan nilai raport Bahasa Indonesia  $\leq 90$  dan nilai raport IPA  $> 80$  maka diterima di sekolah Negeri.

### 3.5 Kebutuhan Pembuatan Sistem

#### 1. 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 :

- a. Processor Intel Core i3
- b. RAM 4 GB
- c. HDD 500 GB
- d. Monitor 12,5"
- e. Mouse

#### 2. Kebutuhan Perangkat Lunak

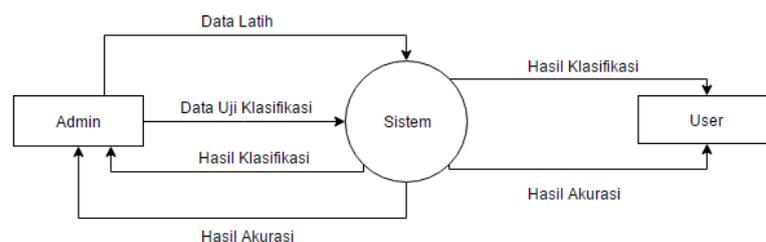
Perangkat lunak adalah program atau aplikasi yang digunakan untuk membangun sistem. Perangkat lunak yang dibutuhkan dalam pembuatan sistem ini adalah :

- a. Sistem Operasi : Windows 10
- b. *Web Server* : XAMPP v3.2.1
- c. *Database Server* : MySQL
- d. Bahasa Pemrograman : PHP
- e. *Software Development* : NetBeans IDE 8.0
- f. *Tool Basis Data* : Sqlyog
- g. *Web Browser* : Google Chrome

### 3.6 Perancangan Sistem

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

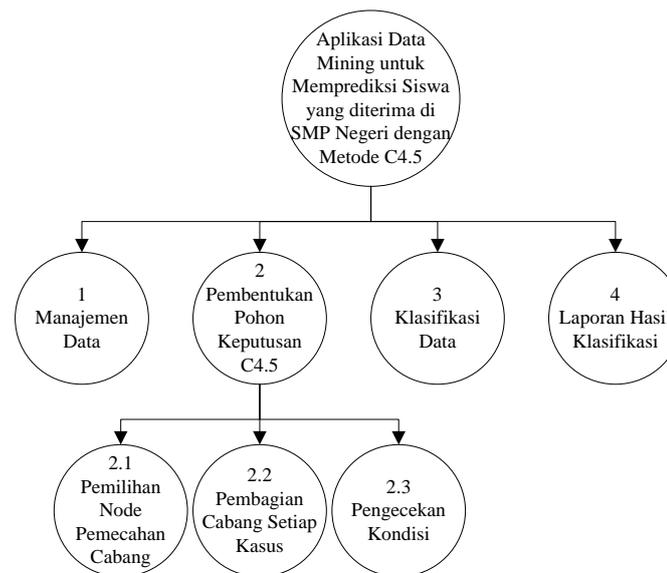
#### 3.5.1 Context Diagram



**Gambar 3.11** *Context Diagram*

Penjelasan dari gambar 3.12, terlihat bahwa yang terlibat (*entity*) dalam sistem ini adalah Admin dan User. Admin, memasukkan data siswa berupa nilai raport semester 1 untuk kelas 6 (enam), nilai try out dan prestasi akademik 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 klasifikasi prediksi siswa berupa kelas siswa masuk negeri atau masuk swasta berdasarkan data yang telah dimasukkan serta hasil akurasi penghitungan. Sedangkan Admin dapat melihat laporan atau daftar hasil klasifikasi prediksi siswa yang masuk sekolah negeri atau swasta yang telah melalui proses klasifikasi.

### 3.5.2 Diagram Berjenjang



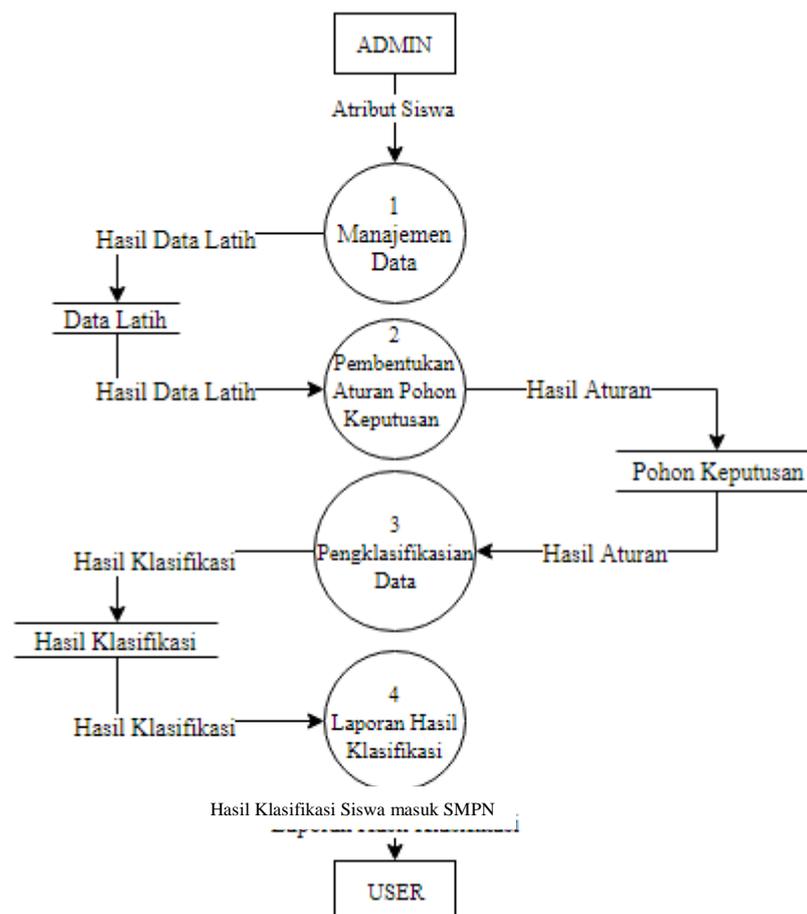
**Gambar 3.12** Diagram Berjenjang

Diagram berjenjang disajikan pada gambar 3.19. berikut penjelasannya:

1. *Top level* : Aplikasi Data Mining untuk Prediksi Siswa yang Masuk SMP Negeri Metode *Decisison Tree C4.5*
2. *Level 0* : 1 Manajemen data, merupakan proses pengolahan data training atau data 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.

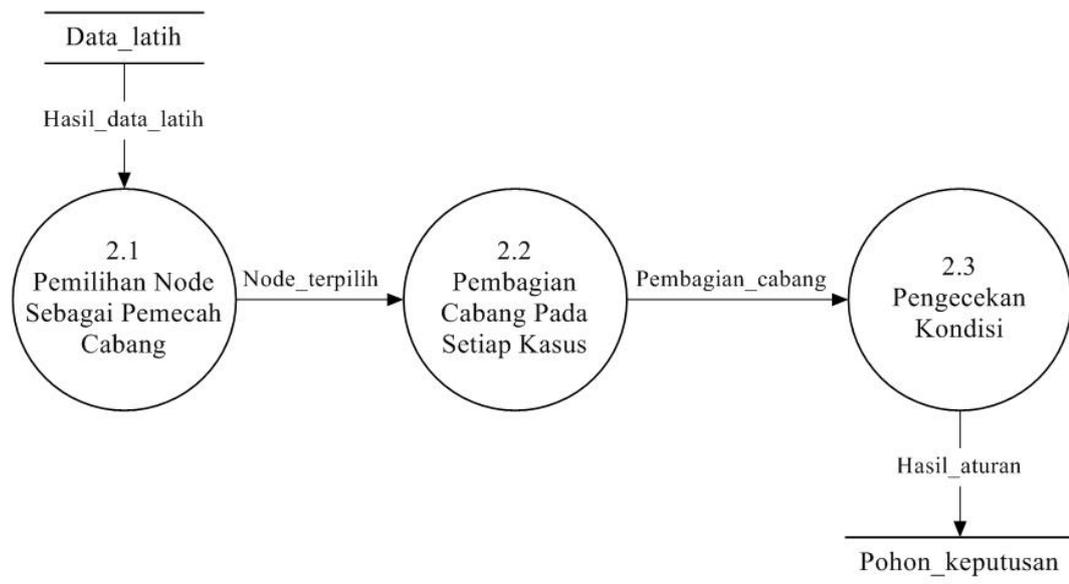
### 3.5.3 Data Flow Diagram Level 0



Gambar 3.13 DFD Level 0

DFD *level 0* pada gambar 3.20 menjelaskan aliran data pada sistem. Terdapat empat proses didalam sistem tersebut. Proses satu adalah manajemen data yang diinputkan oleh admin. Data atribut nilai semester siswa dan prestasi akademik akan menjadi data latih untuk proses pembentukan pohon keputusan. Proses dua adalah pembentukan aturan (pohon keputusan) yang akan digunakan pada proses pengklasifikasian data uji. Hasilnya akan diberikan kepada admin dan akan disimpan dalam hasil klasifikasi. Proses empat adalah pembuatan laporan hasil klasifikasi prediksi siswa yang masuk SMP Negeri yang akan diberikan kepada user dengan mengambil data dari tabel hasil klasifikasi.

### 3.5.4 Data Flow Diagram Level 1



**Gambar 3.14** DFD *Level 1* proses pembentukan aturan (pohon keputusan)

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 *data base*.

### 3.5.5 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 latihan atau data yang akan diproses pada pembentukan pohon keputusan. Tabel ini berisi nilai ujian nasional, nilai *try out* prestasi akademik dan prediksi . Struktur dari tabel ini dapat dilihat pada tabel

**Tabel 3.25** Struktur tabel data latihan

No	Field_name	Type	Length	Key
1	id	Int	11	Primary key
2	Nilai_ipa1	Int	11	
3	Nilai_bin1	Int	11	
4	Nilai_mtk1	Int	11	
5	Nilai_ipa2	Int	11	
6	Nilai_bin2	Int	11	
7	Nilai_mtk2	Int	11	
8	Akademik	char	1	
9	Prediksi	char	1	

#### b. data\_uji

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

**Tabel 3.26** Struktur tabel data uji

No	Field_name	Type	Length	Key
1	id	Int	11	Primary key

2	Nilai_ipa1	Int	11	
3	Nilai_bin1	Int	11	
4	Nilai_mtk1	Int	11	
5	Nilai_ipa2	Int	11	
6	Nilai_bin2	Int	11	
7	Nilai_mtk2	Int	11	
8	Akademik	char	1	
9	Prediksi	char	1	
10	diganosa_sistem	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 3.27** Struktur tabel hasil klasifikasi

No	Field_name	Type	Length	Key
1	id	Int	11	Primary key
2	Nilai_ipa1	Int	11	
3	Nilai_bin1	Int	11	
4	Nilai_mtk1	Int	11	
5	Nilai_ipa2	Int	11	
6	Nilai_bin2	Int	11	
7	Nilai_mtk2	Int	11	
8	Akademik	rchar	1	
9	hasil	char	1	

d. gain

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

**Tabel 3.28** Struktur tabel *gain*

No	Name_field	Type	Length	Key
----	------------	------	--------	-----

1	id	Int	11	Primary key
2	node_id	Int	11	
3	atribut	Varchar	40	
4	gain	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 3.172

**Tabel 3.29** 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 3.30** Struktur t\_keputusan

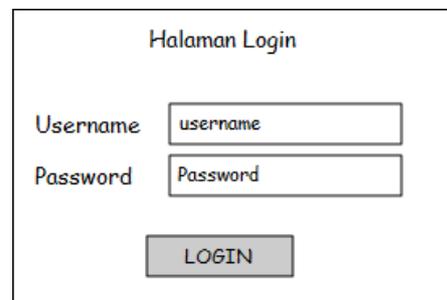
No	Field_name	Type	Length	Key
1	id	Int	11	Primary key
2	parent	Text		
3	akar	Text		
4	keputusan	Varchar	10	

### 3.5.6 Desain Antar Muka

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/i 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 dan kepala sekolah.



The image shows a login form titled "Halaman Login". It contains two input fields: "Username" with the placeholder text "username" and "Password" with the placeholder text "Password". Below the input fields is a button labeled "LOGIN".

**Gambar 3.15** Rancangan halaman *login*

#### b. Halaman Utama

Menu yang ditampilkan untuk pengguna sebagai admin yaitu daftar *user*, Data Latih, dan Klasifikasi C4.5. Sedangkan pengguna sebagai *user* adalah laporan hasil klasifikasi.

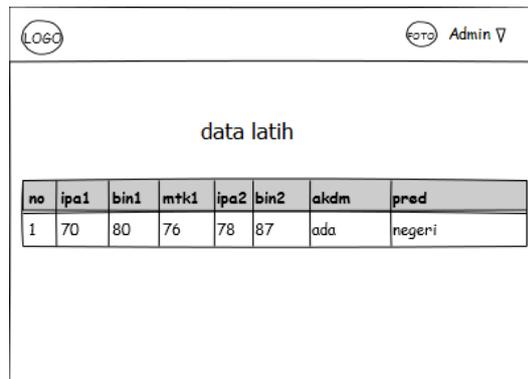


The image shows a dashboard layout. At the top left is a "LOGO" icon, and at the top right is a "foto Admin" icon with a dropdown arrow. The main content area is divided into two columns. The left column is a sidebar menu with the following items: "Dashboard", "Data Latih", "C4.5", "Klasifikasi", "Hasil Klasifikasi", "Prediksi", "Uji Akurasi", and "User". The right column contains the text "Sekolah adalah ....".

**Gambar 3.16** Rancangan halaman utama

### c. Halaman Data Latih

Tampilan ini adalah halaman yang akan digunakan untuk proses pembentukan pohon keputusan. Halaman ini hanya bisa diakses oleh admin saja.

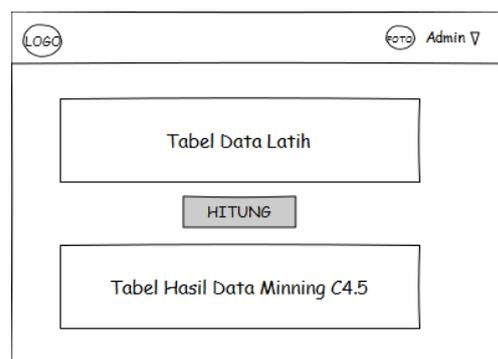


no	ipa1	bin1	mtk1	ipa2	bin2	akdm	pred
1	70	80	76	78	87	ada	negeri

**Gambar 3.17** Rancangan halaman data latih

### d. Halaman Klasifikasi

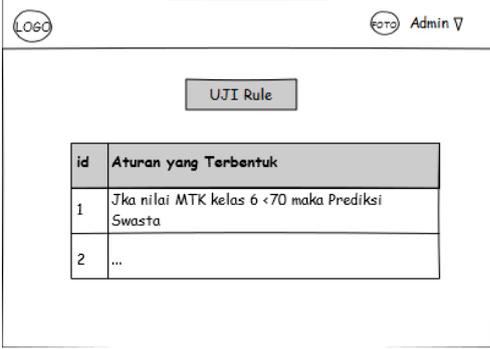
Tampilan ini adalah halaman yang akan digunakan untuk proses pembentukan pohon keputusan.



**Gambar 3.18** Rancangan halaman klasifikasi

e. Halaman Hasil Klasifikasi

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

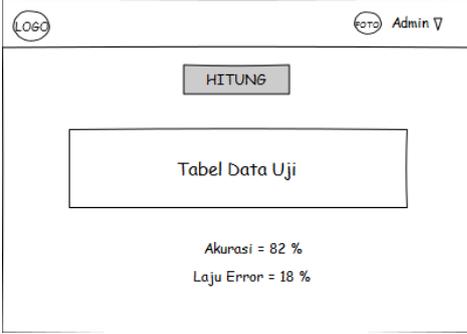


id	Aturan yang Terbentuk
1	Jika nilai MTK kelas 6 <70 maka Prediksi Swasta
2	...

**Gambar 3.19** Rancangan halaman pohon keputusan

f. Halaman Uji Akurasi

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



Akurasi = 82 %  
Laju Error = 18 %

**Gambar 3.20** Rancangan halaman uji pohon keputusan

g. Halaman Prediksi

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

The form for adding a prediction includes a header with a 'LOGO' and 'Admin' profile. The main content area contains several text input fields: 'id', 'nilai ipa 1', 'nilai bin 1', 'nilai mtk 1', 'nilai ipa 2', 'nilai bin 2', 'nilai mtk 2', 'nilai ipa 3', 'nilai bin 3', and 'nilai mtk 3'. Below these fields is a section for 'Prestasi akademik' with two radio buttons labeled 'Ada' and 'Tidak Ada'. At the bottom right, there are two buttons: 'Simpan' and 'batal'.

**Gambar 3.21** Rancangan form penambahan prediksi

#### h. Halaman Ubah Password

Halaman ubah *password* ini adalah halaman untuk mengganti *password* user.

The form for changing a password includes a header with a 'LOGO' and 'Admin' profile. The main content area contains four text input fields: 'id', 'Username', 'Password lama', and 'Password baru'. At the bottom, there are two buttons: 'Simpan' and 'batal'.

**Gambar 3.22** Rancangan halaman ubah *password*

#### i. Halaman Tambah User

Halaman ini digunakan untuk menambahkan user/pengguna untuk sistem ini.

The form for adding a user includes a header with a 'LOGO' and 'Admin' profile. The main content area contains four text input fields: 'id', 'Username', 'Password', and 'Ulangi Password'. At the bottom, there are two buttons: 'Simpan' and 'batal'.

**Gambar 3.23** Rancangan halaman penambahan pengguna

## j. Halaman Daftar User

NO	username	password
1	Admin	****
2	joko	*****

Halaman ini menampilkan daftar user yang ada pada sistem ini.

**Gambar 3.24** Rancangan halaman daftar pengguna

### 3.7 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.

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

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

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.

### 3.8 Skenario Pengujian Sistem

Sebelum membuat aplikasi klasifikasi untuk memprediksi siswa yang masuk dalam SMP Negeri dengan metode decision tree C4.5 ini, perlu

dilakukan beberapa skenario pengujian sistem terlebih dahulu, agar sistem dapat berjalan sesuai dengan tujuan pembuatannya.

- a. Di sediakan 2 macam data yaitu data latih sebanyak 120 dan data uji sebanyak 22. 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 klasifikasi maka sistem dapat menentukan prediksi siswa yang dapat masuk di SMP Negeri dengan kelas “Negeri” dan “Swasta”
- c. Akurasi sistem diperoleh dari hasil prosentase akurasi ketepatan sistem dan hasil prosentase akurasi kesalahan sistem.