

BAB IV

HASIL DAN PEMBAHASAN

4.1 Tahap Pemodelan Data

Pada tahap ini terdapat beberapa proses dalam pemodelan data diantaranya sebagai berikut.

4.1.1 Merancang Design Model Analisis

Pada tahap ini dilakukan proses merancang model analisis yang akan digunakan pada tools Rapid Miner.



Gambar 4. 1 Rancang Model Analisis

4.1.1.1 Data Mentah

Data mentah yang telah dipersiapkan kemudian di import menggunakan tools Rapid miner dengan menggunakan operator Read CSV kemudian dilakukan pemberian label pada data atribut yang akan dijadikan label kelas.



Gambar 4. 2 Operator Read CSV

Import Data - Format your columns.

Format your columns.

Date format: Replace errors with missing values

KE #	CH2O integer	SCC integer	FAF integer	TUE integer	CALC integer	MTRANS integer	NObeyesdad polynomial
1	3	1	2	2	2	4	Obesity_Type_I
2	2	1	2	1	3	1	Obesity_Type_I
3	2	1	0	0	1	1	Obesity_Type_I
4	2	1	0	0	1	1	Obesity_Type_I
5	2	1	0	2	2	4	Obesity_Type_I
6	2	1	0	1	3	1	Obesity_Type_I
7	2	2	0	0	3	1	Obesity_Type_II
8	3	1	0	1	1	4	Obesity_Type_I
9	2	1	0	1	1	4	Obesity_Type_I
10	1	1	2	0	1	4	Obesity_Type_II
11	2	1	0	0	2	1	Obesity_Type_I
12	3	1	3	2	1	5	Obesity_Type_I
13	3	1	2	1	1	4	Obesity_Type_I
14	1	1	3	2	1	5	Obesity_Type_I
15	1	1	2	0	2	1	Obesity_Type_I
16	3	1	0	0	1	2	Obesity_Type_I
17	1	1	1	1	3	1	Obesity_Type_I

no problems.

Previous Finish Cancel

Gambar 4. 3 Data mentah yang telah di import kedalam Tools Rapid Miner

Gambar diatas adalah tampilan dari dataset yang telah di import kedalam aplikasi *rapid miner*

Import Data - Format your columns.

Format your columns.

Date format: Replace errors with missing values

KE #	CH2O integer	SCC integer	FAF integer	TUE integer	CALC integer	MTRANS integer	NObeyesdad polynomial label
1	3	1	2	2	2	4	Obesity_Type_I
2	2	1	2	1	3	1	Obesity_Type_I
3	2	1	0	0	1	1	Obesity_Type_I
4	2	1	0	0	1	1	Obesity_Type_I
5	2	1	0	2	2	4	Obesity_Type_I
6	2	1	0	1	3	1	Obesity_Type_I
7	2	2	0	0	3	1	Obesity_Type_II
8	3	1	0	1	1	4	Obesity_Type_I
9	2	1	0	1	1	4	Obesity_Type_I
10	1	1	2	0	1	4	Obesity_Type_II
11	2	1	0	0	2	1	Obesity_Type_I
12	3	1	3	2	1	5	Obesity_Type_I
13	3	1	2	1	1	4	Obesity_Type_I
14	1	1	3	2	1	5	Obesity_Type_I
15	1	1	2	0	2	1	Obesity_Type_I
16	3	1	0	0	1	2	Obesity_Type_I
17	1	1	1	1	3	1	Obesity_Type_I

no problems.

Previous Finish Cancel

Gambar 4. 4 Data mentah yang telah diberikan label

Gambar diatas adalah data mentah yang telah diberikan fungsi label, dapat dilihat pada kolom warna hijau

Tabel 4. 1 Type Data Attribut Level Obesitas

N0	Attribute	Tipe
1	GENDER	BINOMINAL
2	AGE	INTEGER
3	HEIGHT	INTEGER
4	WEIGHT	INTEGER
5	FAMILY_HISTORY_WITH_OVERWEIGHT	INTEGER
6	FAVC	INTEGER
7	FCVC	INTEGER
8	NCP	INTEGER
9	CAEC	INTEGER
10	SMOKE	INTEGER
11	CH2O	INTEGER
12	SCC	INTEGER
13	FAF	INTEGER
14	TUE	INTEGER
15	CALC	INTEGER
16	MTRANS	INTEGER
17	NOBESITY	POLYNOMINAL

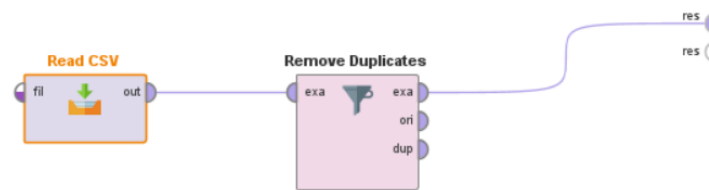
Tabel diatas adalah deksripsi terkait *Attribute* dan *Type* data yang digunakan.

4.1.1.2 Preprocessing Data

Pada tahap ini dilakukan Preprocessing Data sehingga data yang akan dilakukan proses analisis dapat lebih optimal. Berikut ini adalah Tahapan Preprocessing Data :

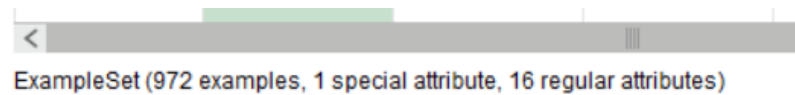
1. Pembersihan data Duplikat

Pada tahap ini menghapus data yang duplikat menggunakan operator *Remove Duplicates*.



Gambar 4. 5 Operator Remove Duplikat

Fungsi Remove Duplikat adalah menghapus Data yang sama didalam dataset



Gambar 4. 6 Sebelum menggunakan Operator Remove duplikat

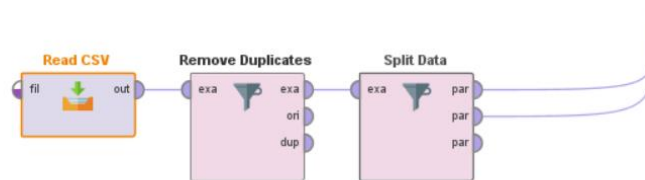
Data sebelum di menggunakan Operator Remove Duplikat dengan jumlah record sebesar 972 record.

< ExampleSet (810 examples, 1 special attribute, 14 regular attributes)

Gambar 4. 7 setelah menggunakan Operator Remove duplikat
 Hasil setelah menggunakan fungsi Remove duplikat maka terlihat data berkurang menjadi 810 data.

4.1.1.3 Data Training

Tahapan selanjutnya adalah melakukan *split data*, operator ini digunakan untuk membagi jumlah data training dan data testing sebesar 70:30.



Gambar 4. 8 Operator Split Data

Operator *Split Data* membagi jumlah data yang akan di uji dan data yang menjadi data *Training*.

Maka pembagiannya akan didapatkan hasil sebagai berikut :

ExampleSet (567 examples, 1 special attribute, 16 regular attributes)

Gambar 4. 9 Data Training

Dari gambar diatas Maka terlihat jumlah data Sample sebesar 567 Data.

ExampleSet (243 examples, 1 special attribute, 16 regular attributes)

Gambar 4. 10 Data Testing

Dan jumlah Data Uji Sebesar 243 Data.

Data Training : 70% x 485 (<i>record dataset</i>)

Data Testing : 30% x 485 (<i>record dataset</i>)
--

4.1.1.4 Model Fit

Pada tahap ini adalah menentukan model Algoritma yang akan digunakan, Untuk penelitian ini menggunakan Algoritma Decision Tree sehingga menghasilkan pohon keputusan. Adapun tahapanya sebagai berikut :

4.1.1.4.1 Menentukan Nilai Entrophy Dan Gain

1. Nilai Entrophy Total

Perhitungan entropy untuk semua data terhadap komposisi kelas, dimana diketahui sebagai berikut :

Tabel 4. 2 Entrophy Total

Attribut	Kriteria	Jumlah Data	Obesity			Entrophy
			Level 1	Level II	Level III	Nilai
NObeyesdad	Label	810	340	268	202	0,980

E (*semua(Total)*)

$$\begin{aligned}
 &= - \left(\left(\frac{340}{810} \right) \times \log_3 \left(\frac{340}{810} \right) \right) + \left(\left(\frac{268}{810} \right) \times \log_3 \left(\frac{268}{810} \right) \right) \\
 &+ \left(\left(\frac{202}{810} \right) \times \log_3 \left(\frac{202}{810} \right) \right) = 0,980
 \end{aligned}$$

2. Nilai Entrophy Dan Gain Atribut

2.1 Atribut Gender (kontinyu).

dimana diketahui data sebagai berikut :

Tabel 4. 3 Atribut Gender

Atribut	Kriteria	Jumlah Data	Obesity		
			Level 1	Level II	Level III
Gender	Male	454	187	266	1
	Female	356	153	2	201
Total		810	340	268	202

E (*semua(Male)*)

$$= - \left(\left(\frac{187}{454} \right) x \log_3 \left(\frac{187}{454} \right) \right) + \left(\left(\frac{266}{454} \right) x \log_3 \left(\frac{266}{454} \right) \right) + \left(\left(\frac{1}{454} \right) x \log_3 \left(\frac{1}{454} \right) \right) = 0.630$$

E (*semua(Female)*)

$$= - \left(\left(\frac{153}{356} \right) x \log_3 \left(\frac{153}{356} \right) \right) + \left(\left(\frac{2}{356} \right) x \log_3 \left(\frac{2}{356} \right) \right) + \left(\left(\frac{201}{356} \right) x \log_3 \left(\frac{201}{356} \right) \right) = 0.651$$

$$Gain$$
 (*semua(Gender)*) = $-\left(\left(\frac{454}{810} \right) x 0.630 \right) + \left(\left(\frac{356}{810} \right) x 0.651 \right) = 0.341$

$$Split$$
 Info (*Gender*) = $-\left(\left(\frac{454}{810} \right) x \log_2 \left(\frac{454}{810} \right) \right) + \left(\left(\frac{356}{810} \right) x \log_2 \left(\frac{356}{810} \right) \right)$

$$= 0.989$$

$$Gain$$
 Rasio (*Gender*) = $\frac{0.341}{0.989} = 0.345$

2.2 Atribut Age (kontinyu)

Dimana diketahui data sebagai berikut :

Tabel 4. 4 Atribut Age

Atribut	Kriteria	Jumlah Data	Obesity		
			Level I	Level II	Level III
Age	<=26	553	235	116	202
	>26	257	105	152	0
Total		810	340	268	202

$$E(\text{semua}(\text{Age} \leq 26))$$

$$= - \left(\left(\frac{235}{553} \right) \times \log_3 \left(\frac{235}{553} \right) \right) + \left(\left(\frac{116}{553} \right) \times \log_3 \left(\frac{116}{553} \right) \right) \\ + \left(\left(\frac{202}{553} \right) \times \log_3 \left(\frac{202}{553} \right) \right) = 0.964$$

$$E(\text{semua}(\text{Age} > 26))$$

$$= - \left(\left(\frac{105}{257} \right) \times \log_3 \left(\frac{105}{257} \right) \right) + \left(\left(\frac{152}{257} \right) \times \log_3 \left(\frac{152}{257} \right) \right) \\ + \left(\left(\frac{0}{257} \right) \times \log_3 \left(\frac{0}{257} \right) \right) = 0.616$$

$$\text{Gain}(\text{semua}(\text{Age})) = - \left(\left(\frac{553}{810} \right) \times 0.964 \right) + \left(\left(\frac{257}{810} \right) \times 0.616 \right) = 0.127$$

$$\text{Split Info}(\text{Age}) = - \left(\left(\frac{553}{810} \right) \times \log_2 \left(\frac{553}{810} \right) \right) + \left(\left(\frac{257}{810} \right) \times \log_2 \left(\frac{257}{810} \right) \right) \\ = 0.901$$

$$\text{Gain Rasio}(\text{Age}) = \frac{0.127}{0.901} = 0.140$$

2.3 Atribut family_history_with_overweight

Dimana diketahui data sebagai berikut :

Tabel 4. 5 Atribut Family History

Atribut	Kriteria	Jumlah Data	Obesity		
			Level 1	Level II	Level III
FH	1	8	7	1	0
	2	802	333	267	202
Total					

$E(\text{semua}(FH\ 1))$

$$= -\left(\left(\frac{7}{8}\right) \times \log_3\left(\frac{7}{8}\right)\right) + \left(\left(\frac{1}{8}\right) \times \log_3\left(\frac{1}{8}\right)\right) + \left(\left(\frac{0}{8}\right) \times \log_3\left(\frac{0}{8}\right)\right)$$

$$= 0,343$$

$E(\text{semua}(FH\ 2))$

$$= -\left(\left(\frac{333}{802}\right) \times \log_3\left(\frac{333}{802}\right)\right) + \left(\left(\frac{267}{802}\right) \times \log_3\left(\frac{267}{802}\right)\right)$$

$$+ \left(\left(\frac{202}{802}\right) \times \log_3\left(\frac{202}{802}\right)\right) = 0.982$$

$$G(\text{semua}(FH)) = -\left(\left(\frac{8}{810}\right) \times 0.343\right) + \left(\left(\frac{802}{810}\right) \times 0.982\right) = 0.005$$

$$\text{Split Info}(FH) = -\left(\left(\frac{8}{810}\right) \times \log_2\left(\frac{8}{810}\right)\right) + \left(\left(\frac{802}{810}\right) \times \log_2\left(\frac{802}{810}\right)\right) = 0.080$$

$$\text{Gain Rasio}(FH) = \frac{0.005}{0.080} = 0.059$$

2.4 Atribut FAVC

Dimana diketahui data sebagai berikut :

Tabel 4. 6 Atribut FAVC

Atribut	Kriteria	Jumlah Data	Obesity		
			Level 1	Level II	Level III
FAVC	1	19	11	7	1
	2	791	329	261	201
Total		810	340	268	202

E (semua(FAVC 1))

$$= -\left(\left(\frac{11}{19}\right) \times \log_3\left(\frac{11}{19}\right)\right) + \left(\left(\frac{7}{19}\right) \times \log_3\left(\frac{7}{19}\right)\right) + \left(\left(\frac{1}{19}\right) \times \log_3\left(\frac{1}{19}\right)\right) = 0,764$$

E (semua(FAVC 2))

$$= -\left(\left(\frac{329}{791}\right) \times \log_3\left(\frac{329}{791}\right)\right) + \left(\left(\frac{261}{791}\right) \times \log_3\left(\frac{261}{791}\right)\right) + \left(\left(\frac{201}{791}\right) \times \log_3\left(\frac{201}{791}\right)\right) = 0.982$$

$$G$$
 (semua(FAVC)) = $-\left(\left(\frac{19}{810}\right) \times 0.764\right) + \left(\left(\frac{791}{810}\right) \times 0.982\right) = 0.003$

$$\text{Split Info (FAVC)} = -\left(\left(\frac{19}{810}\right) \times \log_2\left(\frac{19}{810}\right)\right) + \left(\left(\frac{791}{810}\right) \times \log_2\left(\frac{791}{810}\right)\right) = 0.160$$

$$\text{Gain Rasio (FAVC)} = \frac{0.003}{0.160} = 0.020$$

2.7 Atribut FCVC

Dimana diketahui data sebagai berikut :

Tabel 4. 7 Atribut FCVC

Atribut	Kriteria	Jumlah Data	Obesity		
			Level 1	Level II	Level III
FCVC	1	35	17	18	0
	2	371	247	124	0
	3	404	76	126	202
Total					

E (*semua*(FCVC 1))

$$= - \left(\left(\frac{17}{35} \right) x \log_3 \left(\frac{17}{35} \right) \right) + \left(\left(\frac{18}{35} \right) x \log_3 \left(\frac{18}{35} \right) \right) \\ + \left(\left(\frac{0}{35} \right) x \log_3 \left(\frac{0}{35} \right) \right) = 0,631$$

E (*semua*(FCVC 2))

$$= - \left(\left(\frac{247}{371} \right) x \log_3 \left(\frac{247}{371} \right) \right) + \left(\left(\frac{124}{371} \right) x \log_3 \left(\frac{124}{371} \right) \right) \\ + \left(\left(\frac{0}{371} \right) x \log_3 \left(\frac{0}{371} \right) \right) = 0,580$$

E (*semua*(FCVC 3))

$$= - \left(\left(\frac{76}{404} \right) x \log_3 \left(\frac{76}{404} \right) \right) + \left(\left(\frac{126}{404} \right) x \log_3 \left(\frac{126}{404} \right) \right) \\ + \left(\left(\frac{202}{404} \right) x \log_3 \left(\frac{202}{404} \right) \right) = 0,932$$

Gain (semua(FCVC))

$$= -\left(\left(\frac{35}{810}\right) \times 0.631\right) + \left(\left(\frac{371}{810}\right) \times 0.580\right) + \left(\left(\frac{404}{810}\right) \times 0.932\right)$$

$$= 0.222$$

Split Info (FCVC)

$$= -\left(\left(\frac{35}{810}\right) \times \log_3\left(\frac{35}{810}\right)\right) + \left(\left(\frac{371}{810}\right) \times \log_3\left(\frac{371}{810}\right)\right)$$

$$+ \left(\left(\frac{404}{810}\right) \times \log_3\left(\frac{404}{810}\right)\right) = 0.765$$

$$\text{Gain Rasio (FCVC)} = \frac{0.222}{0.765} = 0.290$$

2.8 Atribut NCP

Dimana diketahui data sebagai berikut :

Tabel 4. 8 Atribut NCP

Atribut	Kriteria	Jumlah Data	Obesity		
			Level I	Level II	Level III
NCP	1	97	74	23	0
	2	76	46	30	0
	3	628	220	206	202
	4	9	0	9	0
Total		810	340	268	202

E (semua(NCP 1))

$$= -\left(\left(\frac{74}{97}\right) \times \log_3\left(\frac{74}{97}\right)\right) + \left(\left(\frac{23}{97}\right) \times \log_3\left(\frac{23}{97}\right)\right)$$

$$+ \left(\left(\frac{0}{97}\right) \times \log_3\left(\frac{0}{97}\right)\right) = 0,499$$

$E(\text{semua}(NCP\ 2))$

$$= -\left(\left(\frac{46}{76}\right) x \log_3 \left(\frac{46}{76}\right)\right) + \left(\left(\frac{30}{76}\right) x \log_3 \left(\frac{30}{76}\right)\right) \\ + \left(\left(\frac{0}{76}\right) x \log_3 \left(\frac{0}{76}\right)\right) = 0.611$$

$E(\text{semua}(NCP\ 3))$

$$= -\left(\left(\frac{220}{628}\right) x \log_3 \left(\frac{220}{628}\right)\right) + \left(\left(\frac{206}{628}\right) x \log_3 \left(\frac{206}{628}\right)\right) \\ + \left(\left(\frac{202}{628}\right) x \log_3 \left(\frac{202}{628}\right)\right) = 0.999$$

$E(\text{semua}(NCP\ 4))$

$$= -\left(\left(\frac{0}{9}\right) x \log_3 \left(\frac{0}{9}\right)\right) + \left(\left(\frac{9}{9}\right) x \log_3 \left(\frac{9}{9}\right)\right) + \left(\left(\frac{0}{9}\right) x \log_3 \left(\frac{0}{9}\right)\right) \\ = 0.000$$

$G(\text{semua}(NCP))$

$$= -\left(\left(\frac{97}{810}\right) x 0.499\right) + \left(\left(\frac{76}{810}\right) x 0.611\right) + \left(\left(\frac{628}{810}\right) x 0.999\right) \\ + \left(\left(\frac{9}{810}\right) x 0.000\right) = 0.088$$

$Split\ Info\ (NCP)$

$$= -\left(\left(\frac{97}{810}\right) x \log_4 \left(\frac{97}{810}\right)\right) + \left(\left(\frac{76}{810}\right) x \log_4 \left(\frac{76}{810}\right)\right) \\ + \left(\left(\frac{628}{810}\right) x \log_4 \left(\frac{628}{810}\right)\right) + \left(\left(\frac{9}{810}\right) x \log_4 \left(\frac{9}{810}\right)\right) = 0.522$$

$$Gain\ Rasio\ (NCP) = \frac{0.088}{0.522} = 0.169$$

2.9 Atribut CAEC

Dimana diketahui data sebagai berikut :

Tabel 4. 9 Atribut CAEC

Atribut	Kriteria	Jumlah Data	Obesity		
			Level 1	Level II	Level III
CAEC	1	2	1	1	0
	2	792	327	264	201
	3	8	6	1	1
	4	8	6	2	0
Total		810	340	268	202

E (semua(CAEC 1))

$$= - \left(\left(\frac{1}{2} \right) x \log_3 \left(\frac{1}{2} \right) \right) + \left(\left(\frac{1}{2} \right) x \log_3 \left(\frac{1}{2} \right) \right) + \left(\left(\frac{0}{2} \right) x \log_3 \left(\frac{0}{2} \right) \right)$$

$$= 0,631$$

E (semua(CAEC 2))

$$= - \left(\left(\frac{327}{792} \right) x \log_3 \left(\frac{327}{792} \right) \right) + \left(\left(\frac{264}{792} \right) x \log_3 \left(\frac{264}{792} \right) \right)$$

$$+ \left(\left(\frac{0}{792} \right) x \log_3 \left(\frac{0}{792} \right) \right) = 0.983$$

E (semua(CAEC 3))

$$= - \left(\left(\frac{6}{8} \right) x \log_3 \left(\frac{6}{8} \right) \right) + \left(\left(\frac{1}{8} \right) x \log_3 \left(\frac{1}{8} \right) \right) + \left(\left(\frac{1}{8} \right) x \log_3 \left(\frac{1}{8} \right) \right)$$

$$= 0.670$$

E (semua(CAEC 4))

$$= - \left(\left(\frac{6}{8} \right) x \log_3 \left(\frac{6}{8} \right) \right) + \left(\left(\frac{2}{8} \right) x \log_3 \left(\frac{2}{8} \right) \right) + \left(\left(\frac{0}{8} \right) x \log_3 \left(\frac{0}{8} \right) \right)$$

$$= 0.512$$

G (semua(CAEC))

$$= -\left(\left(\frac{2}{810}\right) \times 0.631\right) + \left(\left(\frac{792}{810}\right) \times 0.983\right) + \left(\left(\frac{8}{810}\right) \times 0.670\right) \\ + \left(\left(\frac{8}{810}\right) \times 0.512\right) = 0.011$$

$Split Info$ (CAEC)

$$= -\left(\left(\frac{2}{810}\right) \times \log_4\left(\frac{2}{810}\right)\right) + \left(\left(\frac{792}{810}\right) \times \log_4\left(\frac{792}{810}\right)\right) \\ + \left(\left(\frac{8}{810}\right) \times \log_4\left(\frac{8}{810}\right)\right) + \left(\left(\frac{8}{810}\right) \times \log_4\left(\frac{8}{810}\right)\right) = 0.092$$

$$Gain Rasio (NCP) = \frac{0.011}{0.092} = 0.121$$

2.10 Atribut SMOKE

Dimana diketahui data sebagai berikut :

Tabel 4. 10 Atribut Smoke

Atribut	Kriteria	Jumlah Data	Obesity		
			Level 1	Level II	Level III
SMOKE	1	788	334	253	201
	2	22	6	15	1
Total		810	340	268	202

E (semua(SMOKE 1))

$$= -\left(\left(\frac{334}{788}\right) \times \log_3\left(\frac{334}{788}\right)\right) + \left(\left(\frac{253}{788}\right) \times \log_3\left(\frac{253}{788}\right)\right) \\ + \left(\left(\frac{201}{788}\right) \times \log_3\left(\frac{201}{788}\right)\right) = 0,980$$

$$E (\text{semua}(\text{SMOKE } 2))$$

$$= -\left(\left(\frac{6}{371}\right) \times \log_3\left(\frac{6}{22}\right)\right) + \left(\left(\frac{15}{22}\right) \times \log_3\left(\frac{15}{22}\right)\right) \\ + \left(\left(\frac{1}{22}\right) \times \log_3\left(\frac{1}{22}\right)\right) = 0.688$$

$$G (\text{semua}(\text{SMOKE})) = -\left(\left(\frac{788}{810}\right) \times 0.980\right) + \left(\left(\frac{22}{810}\right) \times 0.688\right) = 0.008$$

$$\text{Split Info} (\text{SMOKE}) = -\left(\left(\frac{788}{810}\right) \times \log_2\left(\frac{788}{810}\right)\right) + \left(\left(\frac{22}{810}\right) \times \log_2\left(\frac{22}{810}\right)\right) \\ = 0.180$$

$$\text{Gain Rasio} (\text{SMOKE}) = \frac{0.008}{0.180} = 0.042$$

2.11 Atribut CH20

Dimana diketahui data sebagai berikut :

Tabel 4. 11 Atribut CH20

Atribut	Kriteria	Jumlah Data	Obesity		
			Level 1	Level II	Level III
CH20	1	187	68	74	45
	2	410	165	158	87
	3	213	107	36	70
Total		810	340	268	202

$$E (\text{semua}(\text{CH20 } 1))$$

$$= -\left(\left(\frac{68}{187}\right) \times \log_3\left(\frac{68}{187}\right)\right) + \left(\left(\frac{74}{187}\right) \times \log_3\left(\frac{74}{187}\right)\right) \\ + \left(\left(\frac{45}{187}\right) \times \log_3\left(\frac{45}{187}\right)\right) = 0,981$$

E (*semua*(CH20 2))

$$= -\left(\left(\frac{165}{410}\right) x \log_3 \left(\frac{165}{410}\right)\right) + \left(\left(\frac{158}{410}\right) x \log_3 \left(\frac{158}{410}\right)\right) \\ + \left(\left(\frac{87}{410}\right) x \log_3 \left(\frac{87}{410}\right)\right) = 0.967$$

E (*semua*(CH20 3))

$$= -\left(\left(\frac{107}{213}\right) x \log_3 \left(\frac{107}{213}\right)\right) + \left(\left(\frac{36}{213}\right) x \log_3 \left(\frac{36}{213}\right)\right) \\ + \left(\left(\frac{70}{213}\right) x \log_3 \left(\frac{70}{213}\right)\right) = 0.921$$

G (*semua*(CH20))

$$= -\left(\left(\frac{187}{810}\right) x 0.981\right) + \left(\left(\frac{410}{810}\right) x 0.967\right) + \left(\left(\frac{213}{810}\right) x 0.921\right) \\ = 0.022$$

Split Info (CH20)

$$= -\left(\left(\frac{187}{810}\right) x \log_3 \left(\frac{187}{810}\right)\right) + \left(\left(\frac{410}{810}\right) x \log_3 \left(\frac{410}{810}\right)\right) \\ + \left(\left(\frac{213}{810}\right) x \log_3 \left(\frac{213}{810}\right)\right) = 0.941$$

$$\text{Gain Rasio (CH20)} = \frac{0.022}{0.941} = 0.023$$

2.12 Atribut SCC

Dimana diketahui data sebagai berikut :

Tabel 4. 12 Atribut SCC

Atribut	Kriteria	Jumlah Data	Obesity		
			Level 1	Level II	Level III
SCC	1	807	338	267	202
	2	3	2	1	0
Total		810	340	268	202

E (semua(SCC 1))

$$= - \left(\left(\frac{338}{807} \right) x \log_3 \left(\frac{338}{807} \right) \right) + \left(\left(\frac{267}{807} \right) x \log_3 \left(\frac{267}{807} \right) \right) + \left(\left(\frac{202}{807} \right) x \log_3 \left(\frac{202}{807} \right) \right) = 0,980$$

E (semua(SCC 2))

$$= - \left(\left(\frac{2}{3} \right) x \log_3 \left(\frac{2}{3} \right) \right) + \left(\left(\frac{1}{3} \right) x \log_3 \left(\frac{1}{3} \right) \right) + \left(\left(\frac{0}{3} \right) x \log_3 \left(\frac{0}{3} \right) \right) = 0.579$$

$$G$$
 (semua(SCC)) = $- \left(\left(\frac{807}{810} \right) x 0.980 \right) + \left(\left(\frac{3}{810} \right) x 0.579 \right) = 0.001$

$$Split Info$$
 (SCC) = $- \left(\left(\frac{807}{810} \right) x \log_2 \left(\frac{807}{810} \right) \right) + \left(\left(\frac{3}{810} \right) x \log_2 \left(\frac{3}{810} \right) \right) = 0.035$

$$Gain Rasio$$
 (SCC) = $\frac{0.001}{0.035} = 0.030$

2.13 Atribut FAF

Dimana diketahui data sebagai berikut :

Tabel 4. 13 Atribut FAF

Atribut	Kriteria	Jumlah Data	Obesity		
			Level 1	Level II	Level III
FAF	1	280	127	63	90
	2	321	120	149	52
	3	186	70	56	60
	4	23	23	0	0
Total		810	340	268	202

E (semua(FAF 0))

$$= - \left(\left(\frac{127}{280} \right) x \log_3 \left(\frac{127}{280} \right) \right) + \left(\left(\frac{63}{280} \right) x \log_3 \left(\frac{63}{280} \right) \right) \\ + \left(\left(\frac{90}{280} \right) x \log_3 \left(\frac{90}{280} \right) \right) = 0,964$$

E (semua(FAF 1))

$$= - \left(\left(\frac{120}{321} \right) x \log_3 \left(\frac{120}{321} \right) \right) + \left(\left(\frac{149}{321} \right) x \log_3 \left(\frac{149}{321} \right) \right) \\ + \left(\left(\frac{52}{321} \right) x \log_3 \left(\frac{52}{321} \right) \right) = 0.927$$

E (semua(FAF 2))

$$= - \left(\left(\frac{70}{186} \right) x \log_3 \left(\frac{70}{186} \right) \right) + \left(\left(\frac{56}{186} \right) x \log_3 \left(\frac{56}{186} \right) \right) \\ + \left(\left(\frac{60}{186} \right) x \log_3 \left(\frac{60}{186} \right) \right) = 0.996$$

E (semua(FAF 3))

$$= - \left(\left(\frac{23}{23} \right) x \log_3 \left(\frac{23}{23} \right) \right) + \left(\left(\frac{0}{23} \right) x \log_3 \left(\frac{0}{23} \right) \right) \\ + \left(\left(\frac{0}{23} \right) x \log_3 \left(\frac{0}{23} \right) \right) = 0.000$$

G (semua(FAF))

$$= - \left(\left(\frac{280}{810} \right) x 0.964 \right) + \left(\left(\frac{321}{810} \right) x 0.927 \right) + \left(\left(\frac{186}{810} \right) x 0.996 \right) \\ + \left(\left(\frac{23}{810} \right) x 0.000 \right) = 0.051$$

Split Info (FAF)

$$= - \left(\left(\frac{280}{810} \right) x \log_4 \left(\frac{280}{810} \right) \right) + \left(\left(\frac{321}{810} \right) x \log_4 \left(\frac{321}{810} \right) \right) \\ + \left(\left(\frac{186}{810} \right) x \log_4 \left(\frac{186}{810} \right) \right) + \left(\left(\frac{23}{810} \right) x \log_4 \left(\frac{23}{810} \right) \right) = 0.846$$

$$\text{Gain Rasio (FAF)} = \frac{0.051}{0.846} = 0.060$$

Maka didapatkan nilai *Gain Rasio* untuk Atribut FAF sebesar 0.060

2.14 Atribut TUE

Dimana diketahui data sebagai berikut :

Tabel 4. 14 Atribut TUE

Atribut	Kriteria	Jumlah Data	Obesity		
			Level 1	Level II	Level III
TUE	0	359	165	152	42
	1	374	119	95	160
	2	77	56	21	0

Total	810	340	268	202
-------	-----	-----	-----	-----

$E(\text{semua}(TUE\ 0))$

$$= -\left(\left(\frac{165}{359}\right) x \log_3 \left(\frac{165}{359}\right)\right) + \left(\left(\frac{152}{359}\right) x \log_3 \left(\frac{152}{359}\right)\right) \\ + \left(\left(\frac{42}{359}\right) x \log_3 \left(\frac{42}{359}\right)\right) = 0,885$$

$E(\text{semua}(TUE\ 1))$

$$= -\left(\left(\frac{119}{374}\right) x \log_3 \left(\frac{119}{374}\right)\right) + \left(\left(\frac{95}{374}\right) x \log_3 \left(\frac{95}{374}\right)\right) \\ + \left(\left(\frac{160}{374}\right) x \log_3 \left(\frac{160}{374}\right)\right) = 0,979$$

$E(\text{semua}(TUE\ 2))$

$$= -\left(\left(\frac{56}{77}\right) x \log_3 \left(\frac{56}{77}\right)\right) + \left(\left(\frac{21}{77}\right) x \log_3 \left(\frac{21}{77}\right)\right) \\ + \left(\left(\frac{0}{77}\right) x \log_3 \left(\frac{0}{77}\right)\right) = 0,533$$

$$G(\text{semua}(TUE)) = -\left(\left(\frac{359}{810}\right) x 0,885\right) + \left(\left(\frac{374}{810}\right) x 0,979\right) + \left(\left(\frac{77}{810}\right) x 0,533\right) \\ = 0,085$$

$\text{Split Info}(TUE)$

$$= -\left(\left(\frac{359}{810}\right) x \log_3 \left(\frac{359}{810}\right)\right) + \left(\left(\frac{374}{810}\right) x \log_3 \left(\frac{374}{810}\right)\right) \\ + \left(\left(\frac{77}{810}\right) x \log_3 \left(\frac{77}{810}\right)\right) = 0,857$$

$$\text{Gain Rasio (TUE)} = \frac{0.085}{0.857} = 0.099$$

Maka didapatkan nilai *Gain Rasio* untuk Atribut TUE sebesar 0.099

2.15 Atribut CALC

Dimana diketahui data sebagai berikut :

Tabel 4. 15 Atribut CALC

Atribut	Kriteria	Jumlah Data	Obesity		
			Level 1	Level II	Level III
CALC	1	222	156	65	1
	2	572	170	201	201
	3	16	14	2	0
Total		810	340	268	202

$E(\text{semua(CALC 1)})$

$$= - \left(\left(\frac{156}{222} \right) x \log_3 \left(\frac{156}{222} \right) \right) + \left(\left(\frac{65}{222} \right) x \log_3 \left(\frac{65}{222} \right) \right) + \left(\left(\frac{1}{222} \right) x \log_3 \left(\frac{1}{222} \right) \right) = 0,575$$

$E(\text{semua(CALC 2)})$

$$= - \left(\left(\frac{170}{572} \right) x \log_3 \left(\frac{170}{572} \right) \right) + \left(\left(\frac{201}{572} \right) x \log_3 \left(\frac{201}{572} \right) \right) + \left(\left(\frac{201}{572} \right) x \log_3 \left(\frac{201}{572} \right) \right) = 0.997$$

$E(\text{semua(CALC 3)})$

$$= - \left(\left(\frac{14}{16} \right) x \log_3 \left(\frac{14}{16} \right) \right) + \left(\left(\frac{2}{16} \right) x \log_3 \left(\frac{2}{16} \right) \right) + \left(\left(\frac{0}{16} \right) x \log_3 \left(\frac{0}{16} \right) \right) = 0.343$$

G (*semua*(CALC))

$$= - \left(\left(\frac{222}{810} \right) x 0.575 \right) + \left(\left(\frac{572}{810} \right) x 0.997 \right) + \left(\left(\frac{16}{810} \right) x 0.343 \right)$$

$$= 0.111$$

Split Info (CALC)

$$= - \left(\left(\frac{222}{810} \right) x \log_3 \left(\frac{222}{810} \right) \right) + \left(\left(\frac{572}{810} \right) x \log_3 \left(\frac{572}{810} \right) \right)$$

$$+ \left(\left(\frac{16}{810} \right) x \log_3 \left(\frac{16}{810} \right) \right) = 0.617$$

$$\text{Gain Rasio (CALC)} = \frac{0.111}{0.617} = 0.180$$

Maka didapatkan nilai *Gain Rasio* untuk Atribut CALC sebesar 0.180

2.16 Atribut MTRANS

Dimana diketahui data sebagai berikut :

Tabel 4. 16 Atribut MTRANS

Atribut	Kriteria	Jumlah Data	Obesity		
			Level 1	Level II	Level III
MTRANS	1	199	110	88	1
	2	3	3	0	0
	3	1	0	1	0
	4	604	225	178	201
	5	3	2	1	0
Total		810	340	268	202

$E(\text{semua}(MTARNS 1))$

$$= -\left(\left(\frac{110}{199}\right) x \log_3 \left(\frac{110}{199}\right)\right) + \left(\left(\frac{88}{199}\right) x \log_3 \left(\frac{88}{199}\right)\right) \\ + \left(\left(\frac{1}{199}\right) x \log_3 \left(\frac{1}{199}\right)\right) = 0,651$$

$E(\text{semua}(MTRANS 2))$

$$= -\left(\left(\frac{3}{3}\right) x \log_3 \left(\frac{3}{3}\right)\right) + \left(\left(\frac{0}{3}\right) x \log_3 \left(\frac{0}{3}\right)\right) + \left(\left(\frac{0}{3}\right) x \log_3 \left(\frac{0}{3}\right)\right) \\ = 0.000$$

$E(\text{semua}(MTRANS 3))$

$$= -\left(\left(\frac{0}{1}\right) x \log_3 \left(\frac{0}{1}\right)\right) + \left(\left(\frac{1}{1}\right) x \log_3 \left(\frac{1}{1}\right)\right) + \left(\left(\frac{0}{1}\right) x \log_3 \left(\frac{0}{1}\right)\right) \\ = 0.000$$

$E(\text{semua}(MTRANS 4))$

$$= -\left(\left(\frac{225}{604}\right) x \log_3 \left(\frac{225}{604}\right)\right) + \left(\left(\frac{178}{604}\right) x \log_3 \left(\frac{178}{604}\right)\right) \\ + \left(\left(\frac{201}{604}\right) x \log_3 \left(\frac{201}{604}\right)\right) = 0.996$$

$E(\text{semua}(MTRANS 5))$

$$= -\left(\left(\frac{2}{3}\right) x \log_3 \left(\frac{2}{3}\right)\right) + \left(\left(\frac{1}{3}\right) x \log_3 \left(\frac{1}{3}\right)\right) + \left(\left(\frac{0}{3}\right) x \log_3 \left(\frac{0}{3}\right)\right) \\ = 0.579$$

G (semua(MTRANS))

$$= -\left(\left(\frac{199}{810}\right) \times 0.651\right) + \left(\left(\frac{3}{810}\right) \times 0.000\right) + \left(\left(\frac{1}{810}\right) \times 0.000\right) \\ + \left(\left(\frac{604}{810}\right) \times 0.996\right) + \left(\left(\frac{3}{810}\right) \times 0.579\right) = 0.075$$

Split Info (MTRANS)

$$= -\left(\left(\frac{199}{810}\right) \times \log_5\left(\frac{199}{810}\right)\right) + \left(\left(\frac{3}{810}\right) \times \log_5\left(\frac{3}{810}\right)\right) \\ + \left(\left(\frac{1}{810}\right) \times \log_5\left(\frac{1}{810}\right)\right) + \left(\left(\frac{604}{810}\right) \times \log_5\left(\frac{604}{810}\right)\right) \\ + \left(\left(\frac{3}{810}\right) \times \log_5\left(\frac{3}{810}\right)\right) = 0.381$$

$$\text{Gain Rasio (MTRANS)} = \frac{0.075}{0.381} = 0.198$$

Maka didapatkan nilai *Gain Rasio* untuk Atribut MTRANS sebesar 0.198

3. Hasil Entrophy Atribut Dan Gain dari Entrophy Total

Tabel 4. 17 Hasil Entrophy Atribut Dan Gain dari Entrophy Total

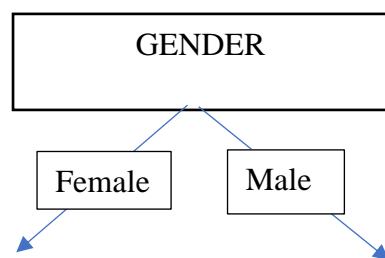
No	Atribut	Value	Entrophy	Gain	Split Info	Gain Rasio
1	Gender	Male	0.630	0.341	0.989	0.345
		Female	0.651			
2	AGE	<=26	0.964	0.127	0.901	0.140
		>26	0.616			
3	FAMILY_HISTORY_WITH_OVERWEIGHT	1	0.343	0.005	0.080	0.059
		2	0.982			
4	FAVC	1	0.764	0.003	0.160	0.020
		2	0.982			
5	FCVC	1	0.631	0.222	0.765	0.290
		2	0.580			

		3	0.932			
6	NCP	1	0.499	0.072	0.522	0.169
		2	0.611			
		3	0.999			
		4	0			
7	CAEC	1	0.631	0.278	0.092	0.121
		2	0.983			
		3	0.670			
		4	0.512			
8	SMOKE	1	0.980	0.008	0.180	0.042
		2	0.688			
9	CH2O	1	0.981	0.022	0.941	0.023
		2	0.967			
		3	0.921			
10	SCC	1	0.980	0.001	0.035	0.030
		2	0.579			
11	FAF	0	0.964	0.051	0.846	0.060
		1	0.927			
		2	0.996			
		3	0			
12	TUE	0	0.885	0.085	0.857	0.099
		1	0.979			
		2	0.533			
13	CALC	1	0.575	0.111	0.617	0.180
		2	0.997			
		3	0.343			
14	MTRANS	1	0.651	0.324	0.381	0.198
		2	0			
		3	0			
		4	0.996			
		5	0.579			

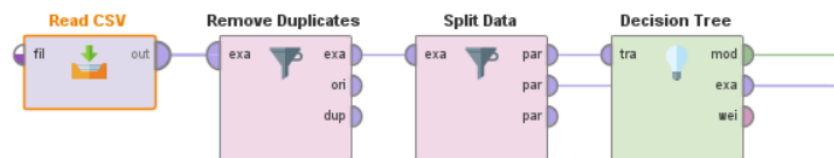
Dari table diatas maka terlihat Gain Rasio yang terbesar maka akan menjadi *Node Root* didalam Pohon Keputusan.

4.1.1.4.2 Menentukan Node Root Pohon Keputusan

Berdasarkan hasil dari perhitungan Manual untuk mencari gain tertinggi dari setiap attribut maka dihasilkan Node root Sebagai berikut :

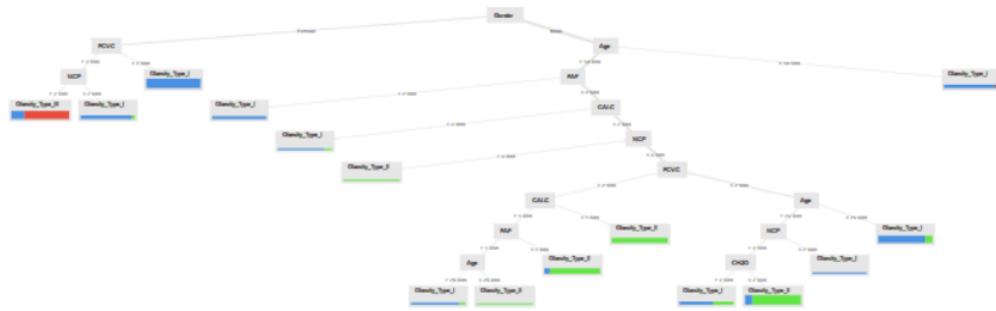


4.1.1.4.3 Menentukan Model Algoritma



Gambar 4. 11 Model Algoritma Decision Tree

Berikut ini hasil Pohon Keputusan dari *Operator Decision Tree* dengan *Node Root* sebagai Gender.



Gambar 4. 12 Pohon Keputusan dari Operator Decision Tree

Terlihat gambar diatas adalah tampilan Susunan Pohon keputusan yang dihasilkan oleh aplikasi *rapid miner*

Tree

```

Gender = Female
| FCVC > 2.500
| | NCP > 2.500: Obesity_Type_III {Obesity_Type_I=10, Obesity_Type_II=0, Obesity_Type_III=34}
| | NCP ≤ 2.500: Obesity_Type_I {Obesity_Type_I=17, Obesity_Type_II=1, Obesity_Type_III=0}
| FCVC ≤ 2.500: Obesity_Type_I {Obesity_Type_I=51, Obesity_Type_II=0, Obesity_Type_III=0}
Gender = Male
| Age > 19.500
| | FAF > 2.500: Obesity_Type_I {Obesity_Type_I=14, Obesity_Type_II=0, Obesity_Type_III=0}
| | FAF ≤ 2.500
| | | CALC > 2.500: Obesity_Type_I {Obesity_Type_I=6, Obesity_Type_II=1, Obesity_Type_III=0}
| | | CALC ≤ 2.500
| | | | NCP > 3.500: Obesity_Type_II {Obesity_Type_I=0, Obesity_Type_II=5, Obesity_Type_III=0}
| | | | NCP ≤ 3.500
| | | | | FCVC > 2.500
| | | | | | CALC > 1.500
| | | | | | | FAF > 1.500
| | | | | | | | Age > 25.500: Obesity_Type_I {Obesity_Type_I=9, Obesity_Type_II=1, Obesity_Type_III=0}
| | | | | | | | Age ≤ 25.500: Obesity_Type_II {Obesity_Type_I=0, Obesity_Type_II=5, Obesity_Type_III=0}
| | | | | | | | FAF ≤ 1.500: Obesity_Type_II {Obesity_Type_I=3, Obesity_Type_II=26, Obesity_Type_III=0}
| | | | | | | CALC ≤ 1.500: Obesity_Type_II {Obesity_Type_I=0, Obesity_Type_II=26, Obesity_Type_III=0}
| | | | | FCVC ≤ 2.500
| | | | | | Age > 23.500
| | | | | | | NCP > 2.500
| | | | | | | | CH20 > 2.500: Obesity_Type_I {Obesity_Type_I=10, Obesity_Type_II=6, Obesity_Type_III=0}
| | | | | | | | CH20 ≤ 2.500: Obesity_Type_II {Obesity_Type_I=7, Obesity_Type_II=46, Obesity_Type_III=0}
| | | | | | | | NCP ≤ 2.500: Obesity_Type_I {Obesity_Type_I=6, Obesity_Type_II=0, Obesity_Type_III=0}
| | | | | | | Age ≤ 23.500: Obesity_Type_I {Obesity_Type_I=34, Obesity_Type_II=5, Obesity_Type_III=0}
| | | | | Age ≤ 19.500: Obesity_Type_I {Obesity_Type_I=16, Obesity_Type_II=0, Obesity_Type_III=1}

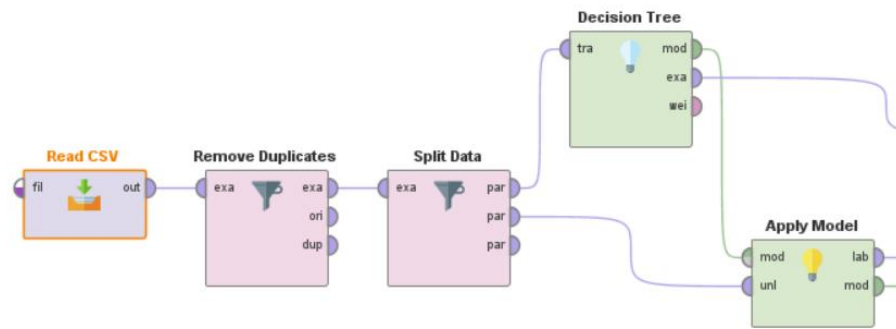
```

Gambar 4. 13 Deskripsi dari pohon keputusan

dapat dilihat dari gambar diatas deksripsi *Attribute* dan nilai dari pohon keputusan.

4.4.1.5 Model Predict

Tahapan selanjutnya adalah menggunakan *Model Predict* untuk melihat prediksi Data Testing berdasarkan Algoritma yang digunakan..



Gambar 4. 14 Model Prediksi untuk uji coba Algoritma

Model prediksi dilakukan menggunakan fungsi Apply Model untuk melihat fungsi algoritma dalam analisis data.

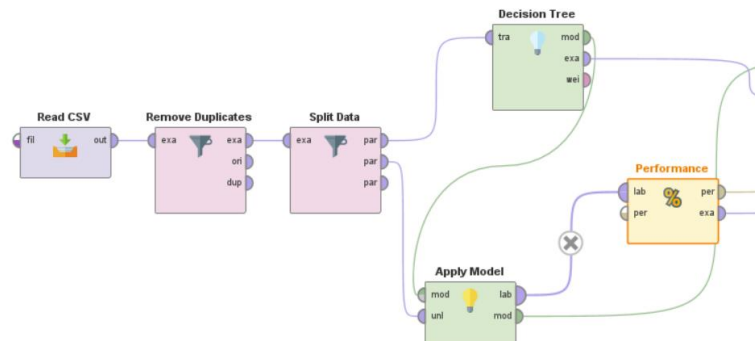
Row No.	NObesyadad	prediction(N...	confidence(Obesity_Type...	confidence(Obesity_Type_II)	confidence(Obesity_Type_III)
1	Obesity_Type_I	Obesity_Type_I	0.900	0.100	0
2	Obesity_Type_I	Obesity_Type_I	0.857	0.143	0
3	Obesity_Type...	Obesity_Type_I	0.857	0.143	0
4	Obesity_Type...	Obesity_Type...	0.227	0	0.773
5	Obesity_Type_I	Obesity_Type...	0.227	0	0.773
6	Obesity_Type_I	Obesity_Type_I	0.872	0.128	0
7	Obesity_Type...	Obesity_Type...	0.132	0.868	0
8	Obesity_Type...	Obesity_Type...	0.132	0.868	0
9	Obesity_Type_I	Obesity_Type_I	0.941	0	0.059
10	Obesity_Type_I	Obesity_Type_I	1	0	0
11	Obesity_Type_I	Obesity_Type_I	0.872	0.128	0
12	Obesity_Type_I	Obesity_Type_I	1	0	0
13	Obesity_Type...	Obesity_Type...	0.103	0.897	0
14	Obesity_Type...	Obesity_Type...	0.227	0	0.773
15	Obesity_Type...	Obesity_Type...	0.227	0	0.773
16	Obesity_Type...	Obesity_Type...	0.227	0	0.773
17	Obesity_Type...	Obesity_Type...	0.227	0	0.773

Gambar 4. 15 Data Testing prediksi. Dari Algoritma DT. C 45.

Dapat dilihat pada gambar diatas, terdapat 2 kolom hijau dan 3 kolom kuning. Dengan kolom hijau pertama sebagai Data yang belum diprediksi dan kolom hijau kedua dari hasil prediksi. Hasil prediksi didapatkan dari value yang ada pada kolom kuning dengan menentukan nilai – nilai pada *attribute*.

4.2 Hasil

4.2.1 Hasil Menggunakan Algoritma C45



Gambar 4. 16 Operator Performa untuk melihat nilai akurasi, Pada tahap ini dilakukan untuk melihat hasil performa model prediksi dari Algoritma Decision Tree.

1. Perhitungan Manual Akurasi

$$Akurasi = \frac{65 + 49 + 15}{65 + 2 + 0 + 8 + 49 + 0 + 5 + 1 + 15} = \frac{129}{142} = 0,889655 \times 100\% = 88,97\%$$

accuracy: 88.97%

	true Obesity_Type_I	true Obesity_Type_II	true Obesity_Type_III	class precision
pred. Obesity_Type_I	65	2	0	97.01%
pred. Obesity_Type_II	8	49	0	85.96%
pred. Obesity_Type_III	5	1	15	71.43%
class recall	83.33%	94.23%	100.00%	

Gambar 4. 17 akurasi dari Algoritma C45

Gambar diatas adalah hasil dari Akurasi menggunakan Algoritma C45 menghasilkan nilai sebesar 88,97%

2. Perhitungan Manual Recall

$$\begin{aligned} \text{Precision Obesity_Type_I} &= \frac{97}{97 + 5 + 0} = \frac{97}{112} = 0.9510 \times 100\% \\ &= 95,10\% \end{aligned}$$

$$\begin{aligned} \text{Precision Obesity_Type_II} &= \frac{78}{78 + 2 + 0} = \frac{78}{80} = 0.9750 \times 100\% \\ &= 97,50\% \end{aligned}$$

$$\text{Precision Obesity_Type_III} = \frac{61}{61 + 0 + 0} = \frac{61}{61} = 1 \times 100\% = 100\%$$

$$\begin{aligned} \text{Precision Total} &= \frac{0.9510 + 0.9750 + 1}{3} = \frac{2.9260}{3} = 0.9753 \times 100\% \\ &= 97,53\% \end{aligned}$$

weighted_mean_recall: 97.53%, weights: 1, 1, 1

	true Obesity_Type_I	true Obesity_Type_II	true Obesity_Type_III	class precision
pred. Obesity_Type_I	97	2	0	97.98%
pred. Obesity_Type_II	5	78	0	93.98%
pred. Obesity_Type_III	0	0	61	100.00%
class recall	95.10%	97.50%	100.00%	

Gambar 4. 18 Gambar. Recall dari Algoritma C45

Gambar diatas adalah hasil dari Recall menggunakan Algoritma C45 menghasilkan nilai sebesar 97,45%

3. Perhitungan Manual Precision

$$\text{Precision Obesity_Type_I} = \frac{97}{97 + 2 + 0} = \frac{97}{99} = 0.9798 \times 100\% = 97,98\%$$

$$\begin{aligned} \text{Precision Obesity_Type_II} &= \frac{78}{78 + 5 + 0} = \frac{78}{83} = 0.9398 \times 100\% \\ &= 93,98\% \end{aligned}$$

$$\text{Precision Obesity_Type_III} = \frac{61}{61 + 0 + 0} = \frac{61}{61} = 1 \times 100\% = 100\%$$

$$Precision\ Total = \frac{0.9798 + 0.9398 + 1}{3} = \frac{2.9196}{3} = 0.9732 \times 100\%$$

$$= 97,32\%$$

weighted_mean_precision: 97.32%, weights: 1, 1, 1

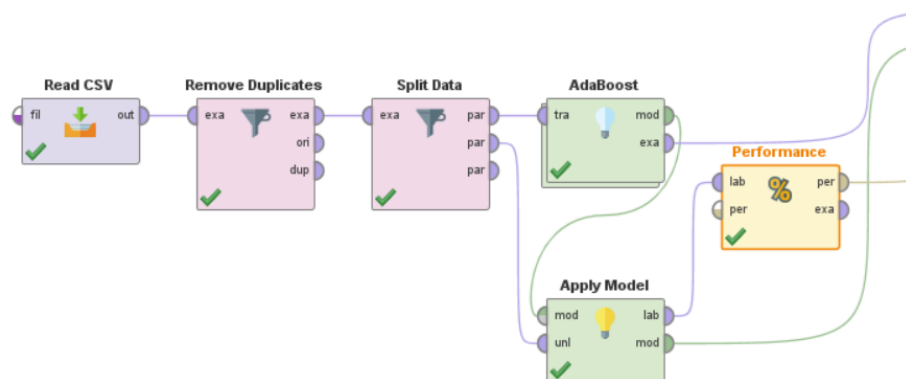
	true Obesity_Type_I	true Obesity_Type_II	true Obesity_Type_III	class precision
pred. Obesity_Type_I	97	2	0	97.98%
pred. Obesity_Type_II	5	78	0	93.98%
pred. Obesity_Type_III	0	0	61	100.00%
class recall	95.10%	97.50%	100.00%	

Gambar 4. 19 Precision dari Algoritma C45

Gambar diatas adalah hasil dari Presisi menggunakan Algoritma C45 menghasilkan nilai sebesar 97,32%

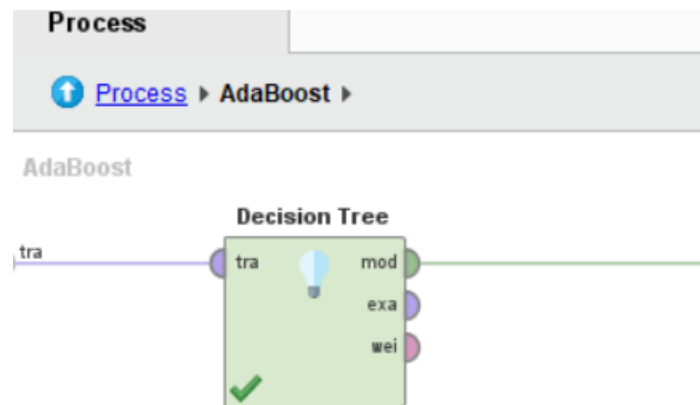
4.2.2 Hasil Menggunakan Algoritma C45 Dan Adaboost

Pada tahapan ini menambahkan Fungsi *Boosting* pada Algoritma untuk memaksimalkan hasil dari Algoritma tersebut.



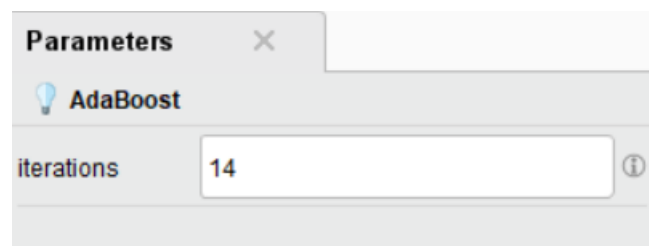
Gambar 4. 20 Model fungsi AdaBoost

Adapun model Boosting menggunakan Adaboost sebagai operatornya dengan nilai *iteration* yang di input sebesar 14.



Gambar 4. 21 Algoritma didalam fungsi AdaBoost

Pada gambar diatas adalah model DT.C45 yang berada di dalam fungsi *Adaboost*



Gambar 4. 22 Parameter iteration

Pada gambar diatas adalah Parameter DT.C45 yang berada di dalam fungsi *Adaboost* dan dilakukan pengeujian maka didapatkan hasil sebagai berikut :

accuracy: 99.59%

	true Obesity_Type_I	true Obesity_Type_II	true Obesity_Type_III	class precision
pred. Obesity_Type_I	102	1	0	99.03%
pred. Obesity_Type_II	0	79	0	100.00%
pred. Obesity_Type_III	0	0	61	100.00%
class recall	100.00%	98.75%	100.00%	

Gambar 4. 23 akurasi dari Algoritma C45 dan Adaboost

Gambar diatas adalah hasil dari akurasi menggunakan Algoritma *C45+Adaboost* menghasilkan nilai sebesar 99,59%

weighted_mean_recall: 99.58%, weights: 1, 1, 1

	true Obesity_Type_I	true Obesity_Type_II	true Obesity_Type_III	class precision
pred. Obesity_Type_I	102	1	0	99.03%
pred. Obesity_Type_II	0	79	0	100.00%
pred. Obesity_Type_III	0	0	61	100.00%
class recall	100.00%	98.75%	100.00%	

Gambar 4. 24 recall dari dari Algoritma C45 dan Adaboost

Gambar diatas adalah\

hasil dari recall menggunakan Algoritma *C45+Adaboost* menghasilkan nilai sebesar 99,598%

weighted_mean_precision: 99.68%, weights: 1, 1, 1

	true Obesity_Type_I	true Obesity_Type_II	true Obesity_Type_III	class precision
pred. Obesity_Type_I	102	1	0	99.03%
pred. Obesity_Type_II	0	79	0	100.00%
pred. Obesity_Type_III	0	0	61	100.00%
class recall	100.00%	98.75%	100.00%	

Gambar 4. 25 Precision dari dari Algoritma C45 dan Adaboost

Gambar diatas adalah hasil dari Presisi menggunakan Algoritma *C45+Adaboost* menghasilkan nilai sebesar 99,68%

4.3 Tahap Evaluasi

4.3.1 Evaluasi Model Algoritma DT.C45

Pada tahap ini dilakukan Evaluasi dari Model Algoritma yang telah digunakan. Dari hasil uji data yang telah dilakukan menggunakan algoritma Decision Tree C45 maka didapatkan nilai sebagai berikut :

No	Keterangan	DT.C45
1	<i>Akurasi</i>	97,12%
2	<i>Recall</i>	97,53%
3	<i>Precision</i>	97,32%

Terlihat pada tabel diatas hasil tersebut sudah sangat baik dengan menghasilkan akurasi, recall dan presisi diatas 90%.

4.3.2 Evaluasi Model Algoritma DT.C45 dan AdaBoost

Pada tahap ini dilakukan Evaluasi dari Model Algoritma yang telah digunakan. Dari hasil uji data yang telah dilakukan menggunakan algoritma Decision Tree C45 dan AdaBoost maka didapatkan nilai sebagai berikut :

No	Keterangan	<i>DT.C45 + Adaboost</i>
1	<i>Akurasi</i>	99,59%
2	<i>Recall</i>	99,58%
3	<i>Precision</i>	99,68%

Hasil DT.C.45 + Adaboost menghasilkan nilai yang lebih baik daripada menggunakan DT.45 diatas 99%.

4.3.3 Evaluasi Perbandingan Model Algoritma DT.C45 dan DT.C45 AdaBoost.

Pada tahap ini dilakukan Evaluasi Perbandingan dari Model Algoritma yang telah digunakan sebagai berikut :

No	Keterangan	<i>DT.C45</i>	<i>DT.C45 + Adaboost</i>
1	<i>Akurasi</i>	97,12%	99,59%
2	<i>Recall</i>	97,53%	99,58%
3	<i>Precision</i>	97,32%	99,68%

Dari data diatas dapat dilihat perbedaan Nilai dengan menggunakan Algoritma *DT.C4* dan *DT.C45* + Ada Boost terlihat Algoritma yang menggunakan *DT.45* + *AdaBoost* menghasilkan nilai yang lebih baik.

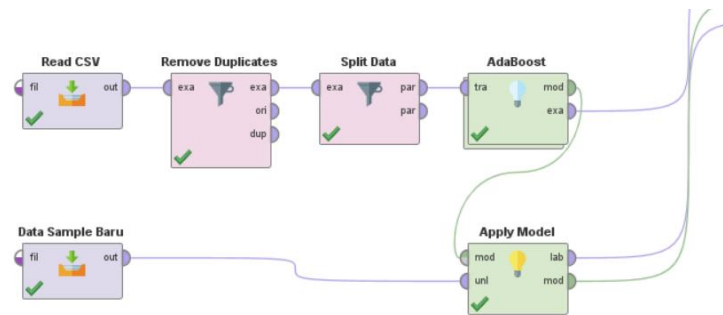
4.3.4 Uji Data Sample Baru

Untuk melihat hasil optimal dari dataset yang telah dianalisis dapat menggunakan data baru sebagai uji coba data.

Terdapat 4 data sample baru berbentuk Csv yang akan di uji coba untuk mengetahui tingkatan obesitasnya.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	F
Gender	Age	Height	Weight	family_his	FAVC	FCVC	NCP	CAEC	SMOKE	CH2O	SCC	FAF	TUE	CALC	MTRANS	NObesdad	
Male	28	185	105	2	2	3	3	3	1	3	1	2	2	2	2	4	
Male	30	176	112	2	2	1	3	1	2	2	2	0	0	0	3	1	
Female	29	153	78	1	2	2	1	2	1	2	1	0	0	0	1	1	
Female	21	175	55	2	2	3	3	2	1	2	1	0	1	2	4		

Gambar 4. 26 Gambar Data baru untuk uji coba.



Gambar 4. 27 Model data uji sample baru

Setelah di uji coba maka didapatkan hasil sebagai berikut :

Row No.	NObeyesdad	prediction(NObeyesdad)	confidence(...)	confidence(...)	confidence(...)	Gender	Age	Height	Weight
1	?	Obesity_Type_I	1	0	0	Male	28	185	105
2	?	Obesity_Type_II	0	1	0	Male	30	176	112
3	?	Obesity_Type_I	1	0	0	Female	29	153	78
4	?	Obesity_Type_I	1	0	0	Female	21	175	55

Gambar 4. 28 Hasil dari uji coba data baru

hasil uji coba prediksi menggunakan *dataset* sebagai *Data Training* dan *Data Prediksi* sebagai *Data uji* yang belum memiliki nilai keputusannya menggunakan aplikasi rapid miner, menghasilkan Prediksi seperti gambar diatas.

4.3.5 Perbandingan Hasil Penelitian

Berdasarkan hasil penelitian yang dilakukan maka didapatkan suatu perbandingan sebagai berikut :

Penelitian	Judul	Akurasi
Mochammad Yusa	EVALUASI MODEL DECISION TREE C4.5 GUNA PREDIKSI POSIBILITAS RESIKO OBESITAS	80%
Wiwik Muslehatin	Penerapan Naïve Bayes Classification untuk Klasifikasi Tingkat Kemungkinan Obesitas Mahasiswa Sistem Informasi UIN Suska Riau	66,67%

Wahyu Hadikristanto	KLASIFIKASI STATUS GIZI ORANG DEWASA MENGUNAKAN ALGORITMA NAÏVE BAYES (STUDI KASUS KLINIK BHAKTI MULIA CIKARANG)	88,67%
Ayu Made Surya Indra Dewi	Implementation Of The K-Nearest Neighbor (KNN) Algorithm For Classification Of Obesity Levels	78,98%
Penelitian Saat ini M. Yhogha Ismail Ibn Ibrahim	PREDIKSI LEVEL OBESITAS MENGGUNAKAN BOOSTING DECISION TREE C45	99,59%

Dari tabel diatas bahwa terjadi peningkatan nilai akurasi pada penelitian saat ini yaitu nilai akurasinya sebesar 99,59%, analisis dari peningkatan nilai akurasi ini jauh lebih baik dibandingkan dengan penelitian sebelumnya karena menerapkan metode dan model analisis yang berbeda pada penerapannya.