

## **BAB III**

### **METODOLOGI PENELITIAN**

Bab ini akan menjelaskan mengenai tahapan metode penelitian yang dilakukan dalam pengerjaan penelitian. Metode penelitian juga digunakan sebagai pedoman dalam pengerjaan penelitian agar mempunyai arah dan terstruktur. Metode yang digunakan dalam penelitian ini adalah metode CRIPS-DM yang meliputi enam fase:

#### **3.1. Fase Pemahaman Bisnis (*Business Understanding*)**

Data pendapatan retribusi Uji KIR periode tahun 2016-2020 menunjukkan antara realisasi dengan target besarnya tidak sama. Selama ini penghitungan target masih dilakukan secara manual. Dari pemahaman bisnis ini perlu adanya metode yang membantu untuk memprediksi pendapatan agar target dan realisasi sesuai.

##### **3.1.1. Tujuan Bisnis**

Tujuan bisnis dilakukan dengan mengenali pola besarnya penerimaan pendapatan untuk mengetahui prediksi pendapatan pada tahun berikutnya.

##### **3.1.2. Situasi Bisnis**

Situasi bisnis yang terjadi pada penelitian ini adalah bahwa besarnya penerimaan pendapatan setiap bulannya tidaklah sama, hal ini dipengaruhi oleh jenis kendaraan dan jumlah kendaraan yang melakukan Uji KIR.

##### **3.1.3. Tujuan Data Mining**

Tujuan data mining ini adalah menggali pengetahuan tentang pola pendapatan dari jenis kendaraan yang berkaitan dengan prediksi pendapatan pada tahun berikutnya.

##### **3.1.4. Perencanaan Strategi**

Strategi awal yang dilakukan adalah pengajuan penelitian ke Dinas Perhubungan Kabupaten Karanganyar, kemudian pengumpulan data dari Dinas tersebut, pengolahan data, terakhir melakukan analisa hasil. Penelitian direncanakan mulai bulan Maret-Juni 2021.

### **3.2. Fase Pemahaman Data (*Data Understanding*)**

Pada tahap ini dilakukan pengumpulan data awal dengan wawancara dan observasi

#### **3.2.1. Wawancara**

Wawancara dilakukan dengan Kepala Dinas Perhubungan Kabupaten Karanganyar yaitu Bapak Sri Suboko, S.Sos., M.Si., dengan mengajukan pertanyaan lisan yang berkaitan dengan kajian yang akan diteliti sebagai landasan dalam memperkuat penelitian dan guna menjawab permasalahan. Teknik wawancara ini digunakan untuk mendapatkan data tentang capaian pendapatan retribusi Uji KIR selama lima tahun periode tahun 2016-2020.

#### **3.2.2. Observasi**

Teknik observasi digunakan untuk melakukan pengamatan secara sistematis terhadap data. Data diambil langsung pada Dinas Perhubungan Kabupaten Karanganyar melalui bendahara pendapatan yaitu Ibu Eni Hastuti. Data yang diminta adalah data pendapatan retribusi Uji KIR tahun 2016-2020, berupa *dataset* "Pendapatan Retribusi Uji KIR.xlsx". Dataset dapat dilihat pada Lampiran 1 Data Pendapatan Retribusi Uji KIR.

### **3.3. Fase Persiapan Data (*Data Preparation*)**

Pada tahap ini dilakukan proses seleksi data yaitu dengan menghapus data yang tidak diperlukan, seperti atribut Total. Hal ini dilakukan karena atribut tersebut tidak berpengaruh terhadap pengolahan data pada proses selanjutnya. Atribut yang digunakan diantaranya: Bulan, Minibus, Mocabus, Bus, Pick Up, Light Truck, Truck, Mobil Baru dan Status. Status naik dan turun diperoleh dari perbandingan data pada bulan sebelumnya. Hasil seleksi data di atas dapat dilihat pada Lampiran 2 Data Seleksi.

### 3.4. Fase Pemodelan (*Modelling*)

Pada tahap ini algoritma yang digunakan adalah algoritma C4.5. Langkah-langkah yang dilakukan untuk menentukan pohon keputusan adalah:

#### 3.4.1. Menyiapkan Data

Data yang akan diolah dalam penelitian ini adalah Pendapatan Retribusi Uji KIR yang sudah diseleksi datanya seperti pada Lampiran 2. Semua atribut bernilai angka dibuat dalam bentuk kategori, penentuan kategori berdasarkan peraturan dari Badan Keuangan Daerah Kabupaten Karanganyar dengan nilai sebagai berikut:

Sangat Kecil = SK : < 10.000.000

Kecil = K : 10.000.000 – 20.000.000

Sedang = S : 20.000.001 – 30.000.000

Besar = B : 30.000.001 – 40.000.000

Sangat Besar = SB : > 40.000.000

Data yang sudah dikategorikan dapat dilihat pada Tabel 3.1 sebagai berikut:

Tabel 3.1 Kategori Data

No	Bulan	Minibus	Microbus	Bus	Pick Up	Light Truck	Truck	Mobil Baru	Status
1	Januari 2016	SK	SK	SK	SB	S	SK	SK	Turun
2	Februari 2016	SK	SK	SK	SB	S	SK	SK	Turun
3	Maret 2016	SK	SK	SK	SB	S	SK	SK	Naik
4	April 2016	SK	SK	SK	SB	S	SK	SK	Turun
5	Mei 2016	SK	SK	SK	SB	B	SK	SK	Naik
6	Juni 2016	SK	SK	SK	SB	S	SK	SK	Naik
7	Juli 2016	SK	SK	SK	SB	S	SK	SK	Turun
8	Agustus 2016	SK	SK	SK	SB	S	SK	SK	Naik
9	September 2016	SK	SK	SK	SB	S	SK	SK	Turun
10	Oktober 2016	SK	SK	SK	SB	S	SK	SK	Naik
11	November 2016	SK	SK	SK	SB	S	SK	SK	Turun
12	Desember 2016	SK	SK	SK	SB	S	SK	SK	Naik
13	Januari 2017	SK	SK	SK	SB	S	SK	SK	Turun
14	Februari 2017	SK	SK	SK	SB	S	SK	SK	Turun
15	Maret 2017	SK	SK	SK	SB	S	SK	SK	Naik
16	April 2017	SK	SK	SK	SB	S	SK	SK	Turun
17	Mei 2017	SK	SK	SK	SB	S	SK	SK	Naik
18	Juni 2017	SK	SK	SK	B	K	SK	SK	Turun
19	Juli 2017	SK	SK	SK	SB	S	SK	SK	Naik
20	Agustus 2017	SK	SK	SK	SB	S	SK	SK	Naik

Lanjutan Tabel 3.1

21	September 2017	SK	SK	SK	S	S	SK	SK	Turun
22	Oktober 2017	SK	SK	SK	SB	S	SK	SK	Naik
23	November 2017	SK	SK	SK	SB	S	SK	SK	Turun
24	Desember 2017	SK	SK	SK	B	S	SK	SK	Turun
25	Januari 2018	SK	SK	SK	SB	S	SK	SK	Naik
26	Februari 2018	SK	SK	SK	SB	S	SK	SK	Turun
27	Maret 2018	SK	SK	SK	SB	S	SK	SK	Naik
28	April 2018	SK	SK	SK	SB	S	SK	SK	Naik
29	Mei 2018	SK	SK	SK	SB	S	SK	SK	Naik
30	Juni 2018	SK	SK	SK	S	K	SK	SK	Turun
31	Juli 2018	SK	SK	SK	SB	S	SK	SK	Naik
32	Agustus 2018	SK	SK	SK	SB	S	SK	SK	Turun
33	September 2018	SK	SK	SK	SB	K	SK	SK	Turun
34	Oktober 2018	SK	SK	SK	SB	S	SK	SK	Naik
35	November 2018	SK	SK	SK	SB	K	SK	SK	Turun
36	Desember 2018	SK	SK	SK	B	K	SK	SK	Turun
37	Januari 2019	SK	SK	SK	SB	S	SK	SK	Naik
38	Februari 2019	SK	SK	SK	SB	K	SK	SK	Turun
39	Maret 2019	SK	SK	SK	SB	S	SK	SK	Naik
40	April 2019	SK	SK	SK	SB	S	SK	SK	Naik
41	Mei 2019	SK	SK	SK	SB	S	SK	SK	Naik
42	Juni 2019	SK	SK	SK	S	K	SK	SK	Turun
43	Juli 2019	SK	SK	SK	SB	S	SK	SK	Naik
44	Agustus 2019	SK	SK	SK	SB	S	SK	SK	Turun
45	September 2019	SK	SK	SK	B	S	SK	K	Naik
46	Oktober 2019	SK	SK	SK	SB	S	SK	SK	Naik
47	November 2019	SK	SK	SK	SB	K	SK	SK	Turun
48	Desember 2019	SK	SK	SK	B	K	SK	SK	Turun
49	Januari 2020	SK	SK	SK	SB	S	SK	SK	Naik
50	Februari 2020	SK	SK	SK	SB	K	SK	SK	Turun
51	Maret 2020	SK	SK	SK	B	S	SK	SK	Turun
52	April 2020	SK	SK	SK	B	K	SK	SK	Turun
53	Mei 2020	SK	SK	SK	S	K	SK	SK	Turun
54	Juni 2020	SK	SK	SK	SB	S	SK	SK	Naik
55	Juli 2020	SK	SK	SK	SB	S	SK	SK	Naik
56	Agustus 2020	SK	SK	SK	SB	K	SK	SK	Turun
57	September 2020	SK	SK	SK	B	S	SK	SK	Turun
58	Oktober 2020	SK	SK	SK	S	K	SK	SK	Turun
59	November 2020	SK	SK	SK	S	K	SK	SK	Naik
60	Desember 2020	SK	SK	SK	B	K	SK	SK	Naik

### 3.4.2. Menghitung Nilai Entropy dan Gain

Dari data Tabel 3.1 diatas diketahui jumlah kasus ada 60, yang berstatus “Naik” 29 *record* dan “Turun” 31 *record*, jumlah kasus dari setiap atribut dapat dilihat pada Tabel 3.2 sebagai berikut:

Tabel 3.2 Jumlah Kasus

		Jml Kasus	Naik	Turun
		S	S <sub>1</sub>	S <sub>2</sub>
Total		60	29	31
Minibus	SK	60	29	31
Microbus	SK	60	29	31
Bus	SK	60	29	31
Pick Up	S	6	1	5
	B	9	2	7
	SB	45	26	19
Light Truck	K	16	2	14
	S	43	26	17
	B	1	1	0
Truck	SK	60	29	31
Mobil Baru	SK	59	28	31
	K	1	1	0

Sehingga didapat nilai *entropy* total:

$$\begin{aligned}
 \text{Entropy Total (S)} &= \sum_{i=1}^n - p_i * \log_2 p_i \\
 &= \left( \frac{-\text{Jml Naik Total}}{\text{Jml Kasus Total}} * \log_2 \left( \frac{\text{Jml Naik Total}}{\text{Jml Kasus Total}} \right) \right) + \\
 &\quad \left( \frac{-\text{Jml Turun Total}}{\text{Jml Kasus Total}} * \log_2 \left( \frac{\text{Jml Turun Total}}{\text{Jml Kasus Total}} \right) \right) \\
 &= \left( \frac{-29}{60} * \log_2 \left( \frac{29}{60} \right) \right) + \left( \frac{-31}{60} * \log_2 \left( \frac{31}{60} \right) \right) \\
 &= 0,50697 + 0,49222 \\
 &= 0,9992
 \end{aligned}$$

Langkah selanjutnya, menghitung nilai *entropy* dari setiap atribut:

$$\begin{aligned}
 \text{Entropy Minibus SK (S)} &= \sum_{i=1}^n - p_i * \log_2 p_i \\
 &= \left( \frac{-\text{Jml Naik Minibus SK}}{\text{Jml Kasus Minibus SK}} * \log_2 \left( \frac{\text{Jml Naik Minibus SK}}{\text{Jml Kasus Minibus SK}} \right) \right) \\
 &+ \left( \frac{-\text{Jml Turun Minibus SK}}{\text{Jml Kasus Minibus SK}} * \log_2 \left( \frac{\text{Jml Turun Minibus SK}}{\text{Jml Kasus Minibus SK}} \right) \right) \\
 &= \left( \frac{-29}{60} * \log_2 \left( \frac{29}{60} \right) \right) + \left( \frac{-31}{60} * \log_2 \left( \frac{31}{60} \right) \right) \\
 &= 0,50697 + 0,49222 \\
 &= 0,9992
 \end{aligned}$$

$$\begin{aligned}
 \text{Entropy Microbus SK (S)} &= \sum_{i=1}^n - p_i * \log_2 p_i \\
 &= \left( \frac{-\text{Jml Naik Microbus SK}}{\text{Jml Kasus Microbus SK}} * \log_2 \left( \frac{\text{Jml Naik Microbus SK}}{\text{Jml Kasus Microbus SK}} \right) \right) \\
 &+ \left( \frac{-\text{Jml Turun Microbus SK}}{\text{Jml Kasus Microbus SK}} * \log_2 \left( \frac{\text{Jml Turun Microbus SK}}{\text{Jml Kasus Microbus SK}} \right) \right) \\
 &= \left( \frac{-29}{60} * \log_2 \left( \frac{29}{60} \right) \right) + \left( \frac{-31}{60} * \log_2 \left( \frac{31}{60} \right) \right) \\
 &= 0,50697 + 0,49222 \\
 &= 0,9992
 \end{aligned}$$

$$\begin{aligned}
 \text{Entropy Bus SK (S)} &= \sum_{i=1}^n - p_i * \log_2 p_i \\
 &= \left( \frac{-\text{Jml Naik Bus SK}}{\text{Jml Kasus Bus SK}} * \log_2 \left( \frac{\text{Jml Naik Bus SK}}{\text{Jml Kasus Bus SK}} \right) \right) + \\
 &\quad \left( \frac{-\text{Jml Turun Bus SK}}{\text{Jml Kasus Bus SK}} * \log_2 \left( \frac{\text{Jml Turun Bus SK}}{\text{Jml Kasus Bus SK}} \right) \right) \\
 &= \left( \frac{-29}{60} * \log_2 \left( \frac{29}{60} \right) \right) + \left( \frac{-31}{60} * \log_2 \left( \frac{31}{60} \right) \right) \\
 &= 0,50697 + 0,49222 \\
 &= 0,9992
 \end{aligned}$$

$$\begin{aligned}
 \text{Entropy Pick Up S (S)} &= \sum_{i=1}^n - p_i * \log_2 p_i \\
 &= \left( \frac{-\text{Jml Naik Pick Up S}}{\text{Jml Kasus Pick Up S}} * \log_2 \left( \frac{\text{Jml Naik Pick Up S}}{\text{Jml Kasus Pick Up S}} \right) \right) + \\
 &\quad \left( \frac{-\text{Jml Turun Pick Up S}}{\text{Jml Kasus Pick Up S}} * \log_2 \left( \frac{\text{Jml Turun Pick Up S}}{\text{Jml Kasus Pick Up S}} \right) \right)
 \end{aligned}$$

$$\begin{aligned}
&= \left( \frac{-1}{6} * \log_2\left(\frac{1}{6}\right) \right) + \left( \frac{-5}{6} * \log_2\left(\frac{5}{6}\right) \right) \\
&= 0,43082 + 0,21919 \\
&= 0,65002
\end{aligned}$$

$$\begin{aligned}
\text{Entropy Pick Up B (S)} &= \sum_{i=1}^n - p_i * \log_2 p_i \\
&= \left( \frac{-\text{Jml Naik Pick Up B}}{\text{Jml Kasus Pick Up B}} * \log_2\left(\frac{\text{Jml Naik Pick Up B}}{\text{Jml Kasus Pick Up B}}\right) \right) + \\
&\quad \left( \frac{-\text{Jml Turun Pick Up B}}{\text{Jml Kasus Pick Up B}} * \log_2\left(\frac{\text{Jml Turun Pick Up B}}{\text{Jml Kasus Pick Up B}}\right) \right) \\
&= \left( \frac{-2}{9} * \log_2\left(\frac{2}{9}\right) \right) + \left( \frac{-7}{9} * \log_2\left(\frac{7}{9}\right) \right) \\
&= 0,48220 + 0,28199 \\
&= 0,7642
\end{aligned}$$

$$\begin{aligned}
\text{Entropy Pick Up SB (S)} &= \sum_{i=1}^n - p_i * \log_2 p_i \\
&= \left( \frac{-\text{Jml Naik Pick Up SB}}{\text{Jml Kasus Pick Up SB}} * \log_2\left(\frac{\text{Jml Naik Pick Up SB}}{\text{Jml Kasus Pick Up SB}}\right) \right) + \\
&\quad \left( \frac{-\text{Jml Turun Pick Up SB}}{\text{Jml Kasus Pick Up SB}} * \log_2\left(\frac{\text{Jml Turun Pick Up SB}}{\text{Jml Kasus Pick Up SB}}\right) \right) \\
&= \left( \frac{-26}{45} * \log_2\left(\frac{26}{45}\right) \right) + \left( \frac{-19}{45} * \log_2\left(\frac{19}{45}\right) \right) \\
&= 0,45726 + 0,52521 \\
&= 0,98247
\end{aligned}$$

$$\begin{aligned}
\text{Entropy Light Truck K (S)} &= \sum_{i=1}^n - p_i * \log_2 p_i \\
&= \left( \frac{-\text{Jml Naik Light Truck K}}{\text{Jml Kasus Light Truck K}} * \log_2\left(\frac{\text{Jml Naik Light Truck K}}{\text{Jml Kasus Light Truck K}}\right) \right) + \\
&\quad \left( \frac{-\text{Jml Turun Light Truck K}}{\text{Jml Kasus Light Truck K}} * \log_2\left(\frac{\text{Jml Turun Light Truck K}}{\text{Jml Kasus Light Truck K}}\right) \right) \\
&= \left( \frac{-2}{16} * \log_2\left(\frac{2}{16}\right) \right) + \left( \frac{-14}{16} * \log_2\left(\frac{14}{16}\right) \right) \\
&= 0,375 + 0,16856 \\
&= 0,54356
\end{aligned}$$

$$\begin{aligned}
\text{Entropy Light Truck S (S)} &= \sum_{i=1}^n - p_i * \log_2 p_i \\
&= \left( \frac{-\text{Jml Naik Light Truck S}}{\text{Jml Kasus Light Truck S}} * \log_2 \left( \frac{\text{Jml Naik Light Truck S}}{\text{Jml Kasus Light Truck S}} \right) \right) + \\
&\quad \left( \frac{-\text{Jml Turun Light Truck S}}{\text{Jml Kasus Light Truck S}} * \log_2 \left( \frac{\text{Jml Turun Light Truck S}}{\text{Jml Kasus Light Truck S}} \right) \right) \\
&= \left( \frac{-26}{43} * \log_2 \left( \frac{26}{43} \right) \right) + \left( \frac{-17}{43} * \log_2 \left( \frac{17}{43} \right) \right) \\
&= 0,43887 + 0,52929 \\
&= 0,96816
\end{aligned}$$

$$\begin{aligned}
\text{Entropy Light Truck B (S)} &= \sum_{i=1}^n - p_i * \log_2 p_i \\
&= \left( \frac{-\text{Jml Naik Light Truck B}}{\text{Jml Kasus Light Truck B}} * \log_2 \left( \frac{\text{Jml Naik Light Truck B}}{\text{Jml Kasus Light Truck B}} \right) \right) + \\
&\quad \left( \frac{-\text{Jml Turun Light Truck B}}{\text{Jml Kasus Light Truck B}} * \log_2 \left( \frac{\text{Jml Turun Light Truck B}}{\text{Jml Kasus Light Truck B}} \right) \right) \\
&= \left( \frac{-1}{1} * \log_2 \left( \frac{1}{1} \right) \right) + \left( \frac{-0}{1} * \log_2 \left( \frac{0}{1} \right) \right) \\
&= 0 + 0 \\
&= 0
\end{aligned}$$

$$\begin{aligned}
\text{Entropy Truck SK (S)} &= \sum_{i=1}^n - p_i * \log_2 p_i \\
&= \left( \frac{-\text{Jml Naik Truck SK}}{\text{Jml Kasus Truck SK}} * \log_2 \left( \frac{\text{Jml Naik Truck SK}}{\text{Jml Kasus Truck SK}} \right) \right) + \\
&\quad \left( \frac{-\text{Jml Turun Truck SK}}{\text{Jml Kasus Truck SK}} * \log_2 \left( \frac{\text{Jml Turun Truck SK}}{\text{Jml Kasus Truck SK}} \right) \right) \\
&= \left( \frac{-29}{60} * \log_2 \left( \frac{29}{60} \right) \right) + \left( \frac{-31}{60} * \log_2 \left( \frac{31}{60} \right) \right) \\
&= 0,50697 + 0,49222 \\
&= 0,9992
\end{aligned}$$

$$\begin{aligned}
\text{Entropy Mobil Baru SK (S)} &= \sum_{i=1}^n - p_i * \log_2 p_i \\
&= \left( \frac{-\text{Jml Naik Mobil Baru SK}}{\text{Jml Kasus Mobil Baru SK}} * \log_2 \left( \frac{\text{Jml Naik Mobil Baru SK}}{\text{Jml Kasus Light Truck SK}} \right) \right) + \\
&\quad \left( \frac{-\text{Jml Turun Mobil Baru SK}}{\text{Jml Kasus Mobil Baru SK}} * \log_2 \left( \frac{\text{Jml Turun Mobil Baru SK}}{\text{Jml Kasus Mobil Baru SK}} \right) \right)
\end{aligned}$$



$$\begin{aligned}
&= \left( \frac{-28}{59} * \log_2\left(\frac{28}{59}\right) \right) + \left( \frac{-31}{59} * \log_2\left(\frac{31}{59}\right) \right) \\
&= 0,51030 + 0,48782 \\
&= 0,99813
\end{aligned}$$

$$\begin{aligned}
Entropy \text{ Mobil Baru K (S)} &= \sum_{i=1}^n - p_i * \log_2 p_i \\
&= \left( \frac{-Jml \text{ Naik Mobil Baru K}}{Jml \text{ Kasus Mobil Baru K}} * \log_2\left(\frac{Jml \text{ Naik Mobil Baru K}}{Jml \text{ Kasus Mobil Baru K}}\right) \right) + \\
&\quad \left( \frac{-Jml \text{ Turun Mobil Baru K}}{Jml \text{ Kasus Mobil Baru K}} * \log_2\left(\frac{Jml \text{ Turun Mobil Baru K}}{Jml \text{ Kasus Mobil Baru K}}\right) \right) \\
&= \left( \frac{-1}{1} * \log_2\left(\frac{1}{1}\right) \right) + \left( \frac{-0}{1} * \log_2\left(\frac{0}{1}\right) \right) \\
&= 0 + 0 \\
&= 0
\end{aligned}$$

Setelah mendapatkan nilai *entropy* dari setiap atribut di atas, berikutnya menghitung nilai *gain*.

$$\begin{aligned}
Gain \text{ Minibus (S,A)} &= Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \\
&= Entropy \text{ Total} - \left( \left( \frac{Jml \text{ Kasus Minibus SK}}{Jml \text{ Kasus Total}} \right) * Entropy \text{ Minibus SK} \right) \\
&= (0,9992) - \left( \left( \frac{60}{60} \right) * 0,9992 \right) \\
&= 0,9992 - 0,9992 \\
&= 0
\end{aligned}$$

$$\begin{aligned}
Gain \text{ Microbus (S,A)} &= Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \\
&= Entropy \text{ Total} - \left( \left( \frac{Jml \text{ Kasus Microbus SK}}{Jml \text{ Kasus Total}} \right) * Entropy \text{ Microbus SK} \right) \\
&= (0,9992) - \left( \left( \frac{60}{60} \right) * 0,9992 \right) \\
&= 0,9992 - 0,9992 \\
&= 0
\end{aligned}$$

$$\begin{aligned}
\text{Gain Bus (S,A)} &= \text{Entropy}(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} \times \text{Entropy}(S_i) \\
&= \text{Entropy Total} - \left( \left( \frac{\text{Jml Kasus Bus SK}}{\text{Jml Kasus Total}} \right) * \text{Entropy Bus SK} \right) \\
&= (0,9992) - \left( \left( \frac{60}{60} \right) * 0,9992 \right) \\
&= 0,9992 - 0,9992 \\
&= 0
\end{aligned}$$

$$\begin{aligned}
\text{Gain Pick Up (S,A)} &= \text{Entropy}(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} \times \text{Entropy}(S_i) \\
&= \text{Entropy Total} - \left( \left( \frac{\text{Jml Kasus Pick Up S}}{\text{Jml Kasus Total}} \right) * \text{Entropy Pick Up S} \right) - \\
&\quad \left( \left( \frac{\text{Jml Kasus Pick Up B}}{\text{Jml Kasus Total}} \right) * \text{Entropy Pick Up B} \right) - \\
&\quad \left( \left( \frac{\text{Jml Kasus Pick Up SB}}{\text{Jml Kasus Total}} \right) * \text{Entropy Pick Up SB} \right) \\
&= (0,9992) - \left( \left( \frac{6}{60} \right) * 0,65002 \right) - \left( \left( \frac{9}{60} \right) * 0,7642 \right) - \left( \left( \frac{45}{60} \right) * 0,98247 \right) \\
&= 0,9992 - 0,06500 - 0,11463 - 0,73685 \\
&= 0,08271
\end{aligned}$$

$$\begin{aligned}
\text{Gain Light Truck (S,A)} &= \text{Entropy}(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} \times \text{Entropy}(S_i) \\
&= \text{Entropy Total} - \left( \left( \frac{\text{Jml Kasus Light Truck K}}{\text{Jml Kasus Total}} \right) * \text{Entropy Light Truck K} \right) - \\
&\quad \left( \left( \frac{\text{Jml Kasus Light Truck S}}{\text{Jml Kasus Total}} \right) * \text{Entropy Light Truck S} \right) - \\
&\quad \left( \left( \frac{\text{Jml Kasus Light Truck B}}{\text{Jml Kasus Total}} \right) * \text{Entropy Light Truck B} \right) \\
&= (0,9992) - \left( \left( \frac{16}{60} \right) * 0,54356 \right) - \left( \left( \frac{43}{60} \right) * 0,96816 \right) - \left( \left( \frac{1}{60} \right) * 0 \right) \\
&= 0,9992 - 0,14495 - 0,69385 - 0 \\
&= 0,160396
\end{aligned}$$

$$\begin{aligned}
\text{Gain Truck (S,A)} &= \text{Entropy}(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} \times \text{Entropy}(S_i) \\
&= \text{Entropy Total} - \left( \left( \frac{\text{Jml Kasus Truck SK}}{\text{Jml Kasus Total}} \right) * \text{Entropy Truck SK} \right) \\
&= (0,9992) - \left( \left( \frac{60}{60} \right) * 0,9992 \right) \\
&= 0,9992 - 0,9992 \\
&= 0
\end{aligned}$$

$$\begin{aligned}
\text{Gain Mobil Baru (S,A)} &= \text{Entropy}(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} \times \text{Entropy}(S_i) \\
&= \text{Entropy Total} - \left( \left( \frac{\text{Jml Kasus Mobil Baru SK}}{\text{Jml Kasus Total}} \right) * \text{Entropy Mobil Baru SK} \right) - \\
&\quad \left( \left( \frac{\text{Jml Kasus Mobil Baru K}}{\text{Jml Kasus Total}} \right) * \text{Entropy Mobil Baru K} \right) \\
&= (0,9992) - \left( \left( \frac{59}{60} \right) * 0,99813 \right) - \left( \left( \frac{1}{60} \right) * 0 \right) \\
&= 0,9992 - 0,98149 - 0 \\
&= 0,0177
\end{aligned}$$

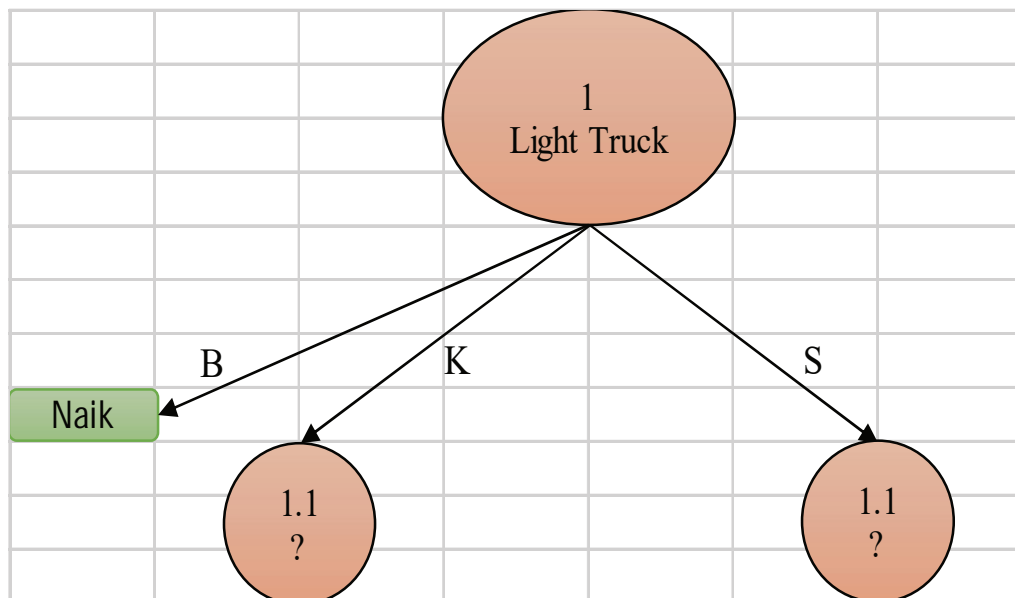
Hasil perhitungan nilai *entropy* dan *gain* dapat dilihat pada Tabel 3.3 sebagai berikut:

Tabel 3.3 Perhitungan Node 1

Node 1		Jml Kasus	Naik	Turun	Entropy	Gain
		S	S <sub>1</sub>	S <sub>2</sub>		
Total		60	29	31	0,999198	
Minibus	SK	60	29	31	0,999198	0
Microbus	SK	60	29	31	0,999198	0
Bus	SK	60	29	31	0,999198	0
Pick Up	S	6	1	5	0,650022	0,08271
	B	9	2	7	0,764205	
	SB	45	26	19	0,982474	
Light Truck	K	16	2	14	0,543564	0,160396
	S	43	26	17	0,968165	
	B	1	1	0	0	
Truck	SK	60	29	31	0,999198	0
Mobil Baru	SK	59	28	31	0,998134	0,0177
	K	1	1	0	0	

### 3.4.3. Menentukan Akar Pohon

Akar dipilih dari atribut yang memiliki nilai gain tertinggi. Pada Tabel 3.3 terlihat gain tertinggi adalah atribut “Light Truck” yaitu 0,160396. Terdapat tiga nilai dari atribut “Light Truck”, yaitu K (Kecil), S (Sedang) dan B (Besar). Dari ketiga nilai tersebut, nilai atribut B (Besar) sudah mengklasifikasikan pada satu kasus yaitu keputusannya “Naik” sehingga tidak perlu dilakukan perhitungan lagi. Untuk nilai atribut K (Kecil) dan S (Sedang) masih perlu dilakukan perhitungan lagi. Dari hasil tersebut dapat digambarkan pohon keputusan seperti pada Gambar 3.1 sebagai berikut:



Gambar 3.1 Pohon Keputusan Node 1

### 3.4.4. Mengulangi Proses

Proses ke-2

Selanjutnya adalah menghitung Node 1.1 sebagai akar cabang, caranya sama dengan proses di atas dengan menghitung nilai entropy dan gain dari atribut yang belum terpilih, yaitu “Minibus”, “Microbus”, “Bus”, “Pick Up”, “Truck” dan “Mobil Baru”.

Berdasarkan Gambar 3.1 atribut “Light Truck:K” digunakan untuk memfilter data, sehingga diperoleh data kasus pada Tabel 3.4 sebagai berikut:

Tabel 3.4 Data dengan filter Light Truck:K

No	Bulan	Mibus	Microbus	Bus	Pick Up	Light Truck	Truck	Mobil Baru	Status
1	Juni 2017	SK	SK	SK	B	K	SK	SK	Turun
2	Juni 2018	SK	SK	SK	S	K	SK	SK	Turun
3	September 2018	SK	SK	SK	SB	K	SK	SK	Turun
4	November 2018	SK	SK	SK	SB	K	SK	SK	Turun
5	Desember 2018	SK	SK	SK	B	K	SK	SK	Turun
6	Februari 2019	SK	SK	SK	SB	K	SK	SK	Turun
7	Juni 2019	SK	SK	SK	S	K	SK	SK	Turun
8	November 2019	SK	SK	SK	SB	K	SK	SK	Turun
9	Desember 2019	SK	SK	SK	B	K	SK	SK	Turun
10	Februari 2020	SK	SK	SK	SB	K	SK	SK	Turun
11	April 2020	SK	SK	SK	B	K	SK	SK	Turun
12	Mei 2020	SK	SK	SK	S	K	SK	SK	Turun
13	Agustus 2020	SK	SK	SK	SB	K	SK	SK	Turun
14	Oktober 2020	SK	SK	SK	S	K	SK	SK	Turun
15	November 2020	SK	SK	SK	S	K	SK	SK	Naik
16	Desember 2020	SK	SK	SK	B	K	SK	SK	Naik

Selanjutnya menghitung:

- a. Jumlah kasus
- b. Jumlah kasus keputusan "Naik"
- c. Jumlah kasus keputusan "Turun"
- d. Nilai entropy dan gain

Hasil perhitungan dapat dilihat pada Tabel 3.5 sebagai berikut:

Tabel 3.5 Perhitungan Node 1.1 Light Truck:K

Node 1.1 Light Truck:K		Jml Kasus	Naik	Turun	Entropy	Gain
		S	S <sub>1</sub>	S <sub>2</sub>		
Total		16	2	14	0,54356	
Minibus	SK	16	2	14	0,54356	0
Microbus	SK	16	2	14	0,54356	0
Bus	SK	16	2	14	0,54356	0
Pick Up	S	5	1	4	0,72193	0,09236
	B	5	1	4	0,72193	
	SB	6	0	6	0	
Truck	SK	16	2	14	0,54356	0
Mobil Baru	SK	16	2	14	0,54356	0

Dari hasil perhitungan Tabel 3.5 didapat gain tertinggi atribut “Pick Up” yaitu 0,09236, maka atribut “Pick Up” terpilih sebagai akar cabang dari atribut “Light Truck” yang bernilai K (Kecil). Atribut “Pick Up” terdapat tiga nilai yaitu S (Sedang), B (Besar) dan SB (Sangat Besar). Karena nilai SB (Sangat Besar) sudah mengklasifikasikan pada satu kasus yaitu keputusannya “Turun” maka tidak perlu dilakukan perhitungan lagi. Sedangkan nilai S (Sedang) dan B (Besar) masih perlu dilakukan perhitungan lagi. Untuk nilai S (Sedang) dan B (Besar) memiliki nilai entropy yang sama, maka dipilih salah satu saja, yaitu dipilih B (Besar).

Berdasarkan Gambar 3.1 atribut “Light Truck:S” digunakan untuk memfilter data, sehingga diperoleh data kasus pada Tabel 3.6 sebagai berikut:

Tabel 3.6 Data dengan filter Light Truck:S

No	Bulan	Minibus	Microbus	Bus	Pick Up	Light Truck	Truck	Mobil Baru	Status
1	Januari 2016	SK	SK	SK	SB	S	SK	SK	Turun
2	Februari 2016	SK	SK	SK	SB	S	SK	SK	Turun
3	Maret 2016	SK	SK	SK	SB	S	SK	SK	Naik
4	April 2016	SK	SK	SK	SB	S	SK	SK	Turun
5	Juni 2016	SK	SK	SK	SB	S	SK	SK	Naik
6	Juli 2016	SK	SK	SK	SB	S	SK	SK	Turun
7	Agustus 2016	SK	SK	SK	SB	S	SK	SK	Naik
8	September 2016	SK	SK	SK	SB	S	SK	SK	Turun
9	Oktober 2016	SK	SK	SK	SB	S	SK	SK	Naik
10	November 2016	SK	SK	SK	SB	S	SK	SK	Turun
11	Desember 2016	SK	SK	SK	SB	S	SK	SK	Naik
12	Januari 2017	SK	SK	SK	SB	S	SK	SK	Turun
13	Februari 2017	SK	SK	SK	SB	S	SK	SK	Turun
14	Maret 2017	SK	SK	SK	SB	S	SK	SK	Naik
15	April 2017	SK	SK	SK	SB	S	SK	SK	Turun
16	Mei 2017	SK	SK	SK	SB	S	SK	SK	Naik
17	Juli 2017	SK	SK	SK	SB	S	SK	SK	Naik
18	Agustus 2017	SK	SK	SK	SB	S	SK	SK	Naik
19	September 2017	SK	SK	SK	S	S	SK	SK	Turun
20	Oktober 2017	SK	SK	SK	SB	S	SK	SK	Naik
21	November 2017	SK	SK	SK	SB	S	SK	SK	Turun
22	Desember 2017	SK	SK	SK	B	S	SK	SK	Turun
23	Januari 2018	SK	SK	SK	SB	S	SK	SK	Naik
24	Februari 2018	SK	SK	SK	SB	S	SK	SK	Turun
25	Maret 2018	SK	SK	SK	SB	S	SK	SK	Naik
26	April 2018	SK	SK	SK	SB	S	SK	SK	Naik
27	Mei 2018	SK	SK	SK	SB	S	SK	SK	Naik
28	Juli 2018	SK	SK	SK	SB	S	SK	SK	Naik
29	Agustus 2018	SK	SK	SK	SB	S	SK	SK	Turun
30	Oktober 2018	SK	SK	SK	SB	S	SK	SK	Naik
31	Januari 2019	SK	SK	SK	SB	S	SK	SK	Naik
32	Maret 2019	SK	SK	SK	SB	S	SK	SK	Naik
33	April 2019	SK	SK	SK	SB	S	SK	SK	Naik
34	Mei 2019	SK	SK	SK	SB	S	SK	SK	Naik
35	Juli 2019	SK	SK	SK	SB	S	SK	SK	Naik
36	Agustus 2019	SK	SK	SK	SB	S	SK	SK	Turun
37	September 2019	SK	SK	SK	B	S	SK	K	Naik
38	Oktober 2019	SK	SK	SK	SB	S	SK	SK	Naik
39	Januari 2020	SK	SK	SK	SB	S	SK	SK	Naik
40	Maret 2020	SK	SK	SK	B	S	SK	SK	Turun
41	Juni 2020	SK	SK	SK	SB	S	SK	SK	Naik
42	Juli 2020	SK	SK	SK	SB	S	SK	SK	Naik
43	September 2020	SK	SK	SK	B	S	SK	SK	Turun

Selanjutnya menghitung:

- Jumlah kasus
- Jumlah kasus keputusan “Naik”
- Jumlah kasus keputusan “Turun”
- Nilai entropy dan gain

Hasil perhitungan dapat dilihat pada Tabel 3.7 sebagai berikut:

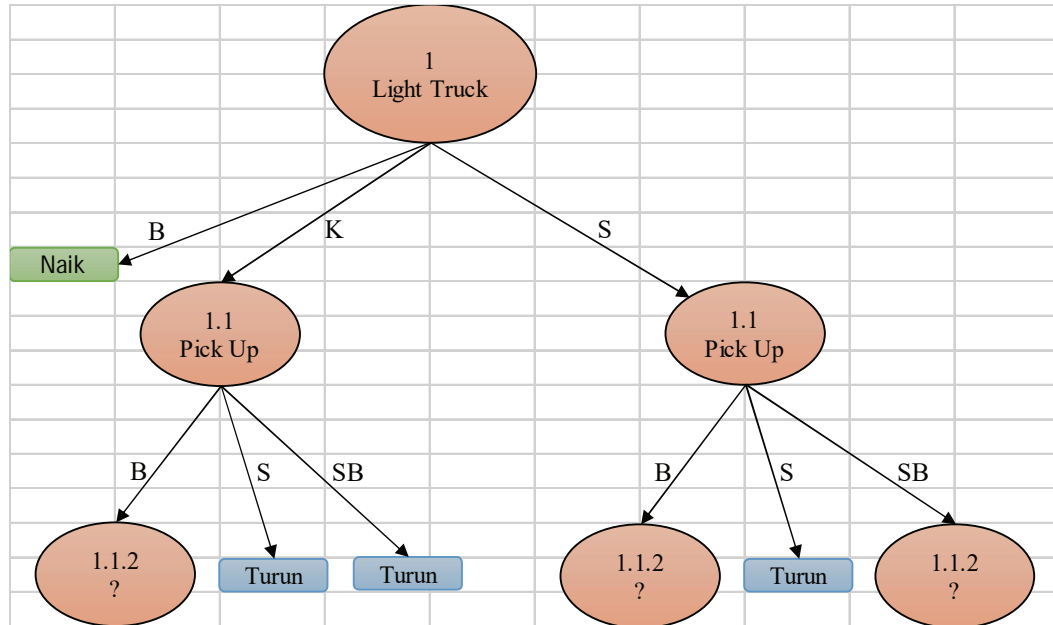
Tabel 3.7 Perhitungan Node 1.1 Light Truck:S

Node 1.1 Light Truck:S		Jml Kasus	Naik	Turun	Entropy	Gain
		S	S <sub>1</sub>	S <sub>2</sub>		
Total		43	26	17	0,96816	
Minibus	SK	43	26	17	0,96816	0
Microbus	SK	43	26	17	0,96816	0
Bus	SK	43	26	17	0,96816	0
Pick Up	S	1	0	1	0	0,07365
	B	4	1	3	0,81128	
	SB	38	25	13	0,92682	
Truck	SK	43	26	17	0,96816	0
Mobil Baru	SK	42	25	17	0,97367	0,01714
	K	1	1	0	0	

Dari hasil perhitungan Tabel 3.7 didapat gain tertinggi atribut “Pick Up” yaitu 0,07365, maka atribut “Pick Up” terpilih sebagai akar cabang dari atribut “Light Truck” yang bernilai S (Sedang). Atribut “Pick Up” terdapat tiga nilai yaitu S (Sedang), B (Besar) dan SB (Sangat Besar). Karena nilai S (Sedang) sudah mengklasifikasikan pada satu kasus yaitu keputusannya “Turun” maka tidak perlu dilakukan perhitungan lagi. Sedangkan nilai B (Besar) dan SB (Sangat Besar) masih perlu dilakukan perhitungan lagi.



Dari hasil proses ke-2 di atas dapat digambarkan pohon keputusan seperti pada Gambar 3.2 sebagai berikut:



Gambar 3.2 Pohon Keputusan Node 1.1

Proses ke-3

Selanjutnya adalah menghitung Node 1.1.2 sebagai akar cabang, perhitungan dilakukan dengan cara yang sama pada proses ke-2 dengan menghitung nilai entropy dan gain dari atribut yang belum terpilih, yaitu “Minibus”, “Microbus”, “Bus”, “Truck” dan “Mobil Baru”.

Berdasarkan Gambar 3.2 kita gunakan atribut “Light Truck:K dan Pick Up:B” untuk memfilter data, sehingga diperoleh data kasus pada Tabel 3.8 sebagai berikut:

Tabel 3.8 Data dengan filter Light Truck:K dan Pick Up:B

No	Bulan	Minibus	Microbus	Bus	Pick Up	Light Truck	Truck	Mobil Baru	Status
1	Juni 2017	SK	SK	SK	B	K	SK	SK	Turun
2	Desember 2018	SK	SK	SK	B	K	SK	SK	Turun
3	Desember 2019	SK	SK	SK	B	K	SK	SK	Turun
4	April 2020	SK	SK	SK	B	K	SK	SK	Turun
5	Desember 2020	SK	SK	SK	B	K	SK	SK	Naik

Selanjutnya menghitung:

- Jumlah kasus
- Jumlah kasus keputusan “Naik”
- Jumlah kasus keputusan “Turun”
- Nilai entropy dan gain

Hasil perhitungan dapat dilihat pada Tabel 3.9 sebagai berikut:

Tabel 3.9 Perhitungan Node 1.1.2 Light Truck:K dan Pick Up:B

Node 1.1.2 Light Truck:K dan Pick Up:B		Jml Kasus	Naik	Turun	Entropy	Gain
		S	S <sub>1</sub>	S <sub>2</sub>		
Total		5	1	4	0,72193	
Minibus	SK	5	1	4	0,72193	0
Microbus	SK	5	1	4	0,72193	0
Bus	SK	5	1	4	0,72193	0
Truck	SK	5	1	4	0,72193	0
Mobil Baru	SK	5	1	4	0,72193	0

Dari hasil perhitungan Tabel 3.9 terlihat kelima nilai atribut sudah mengklasifikasikan pada satu kasus yaitu keputusannya “Turun” maka tidak perlu dilakukan perhitungan lagi.

Berdasarkan Gambar 3.2 kita gunakan atribut “Light Truck:S dan Pick Up:B” untuk memfilter data, sehingga diperoleh data kasus pada Tabel 3.10 sebagai berikut:

Tabel 3.10 Data dengan filter Light Truck:S dan Pick Up:B

No	Bulan	Minibus	Microbus	Bus	Pick Up	Light Truck	Truck	Mobil Baru	Status
1	Desember 2017	SK	SK	SK	B	S	SK	SK	Turun
2	September 2019	SK	SK	SK	B	S	SK	K	Naik
3	Maret 2020	SK	SK	SK	B	S	SK	SK	Turun
4	September 2020	SK	SK	SK	B	S	SK	SK	Turun

Selanjutnya menghitung:

- Jumlah kasus
- Jumlah kasus keputusan “Naik”
- Jumlah kasus keputusan “Turun”
- Nilai entropy dan gain

Hasil perhitungan dapat dilihat pada Tabel 3.11 sebagai berikut:

Tabel 3.11 Perhitungan Node 1.1.2 Light Truck:S dan Pick Up:B

Node 1.1.2 Light Truck:S dan Pick Up:B		Jml Kasus	Naik	Turun	Entropy	Gain
		S	S <sub>1</sub>	S <sub>2</sub>		
Total		4	1	3	0,81128	
Minibus	SK	4	1	3	0,81128	0
Microbus	SK	4	1	3	0,81128	0
Bus	SK	4	1	3	0,81128	0
Truck	SK	4	1	3	0,81128	0
Mobil Baru	SK	3	0	3	0	0,81128
	K	1	1	0	0	

Dari hasil perhitungan Tabel 3.11 didapat gain tertinggi atribut “Mobil Baru” yaitu 0,81128, maka atribut “Mobil Baru” terpilih sebagai akar cabang dari atribut “Pick Up” yang bernilai B (Besar). Atribut “Mobil Baru” terdapat dua nilai yaitu SK (Sangat Kecil) dan K (Kecil). Karena nilai SK (Sangat Kecil) sudah mengklasifikasikan pada satu kasus yaitu keputusannya “Turun” dan nilai K (Kecil) sudah mengklasifikasikan pada satu kasus yaitu keputusannya “Naik” maka tidak perlu dilakukan perhitungan lagi.

Berdasarkan Gambar 3.2 kita gunakan atribut “Light Truck:S dan Pick Up:SB” untuk memfilter data, sehingga diperoleh data kasus pada Tabel 3.12 sebagai berikut:

Tabel 3.12 Data dengan filter Light Truck:S dan Pick Up:SB

No	Bulan	Minibus	Microbus	Bus	Pick Up	Light Truck	Truck	Mobil Baru	Status
1	Januari 2016	SK	SK	SK	SB	S	SK	SK	Turun
2	Februari 2016	SK	SK	SK	SB	S	SK	SK	Turun
3	Maret 2016	SK	SK	SK	SB	S	SK	SK	Naik
4	April 2016	SK	SK	SK	SB	S	SK	SK	Turun
5	Juni 2016	SK	SK	SK	SB	S	SK	SK	Naik
6	Juli 2016	SK	SK	SK	SB	S	SK	SK	Turun
7	Agustus 2016	SK	SK	SK	SB	S	SK	SK	Naik
8	September 2016	SK	SK	SK	SB	S	SK	SK	Turun
9	Oktober 2016	SK	SK	SK	SB	S	SK	SK	Naik
10	November 2016	SK	SK	SK	SB	S	SK	SK	Turun
11	Desember 2016	SK	SK	SK	SB	S	SK	SK	Naik
12	Januari 2017	SK	SK	SK	SB	S	SK	SK	Turun
13	Februari 2017	SK	SK	SK	SB	S	SK	SK	Turun
14	Maret 2017	SK	SK	SK	SB	S	SK	SK	Naik
15	April 2017	SK	SK	SK	SB	S	SK	SK	Turun
16	Mei 2017	SK	SK	SK	SB	S	SK	SK	Naik
17	Juli 2017	SK	SK	SK	SB	S	SK	SK	Naik
18	Agustus 2017	SK	SK	SK	SB	S	SK	SK	Naik
19	Oktober 2017	SK	SK	SK	SB	S	SK	SK	Naik
20	November 2017	SK	SK	SK	SB	S	SK	SK	Turun
21	Januari 2018	SK	SK	SK	SB	S	SK	SK	Naik
22	Februari 2018	SK	SK	SK	SB	S	SK	SK	Turun
23	Maret 2018	SK	SK	SK	SB	S	SK	SK	Naik
24	April 2018	SK	SK	SK	SB	S	SK	SK	Naik
25	Mei 2018	SK	SK	SK	SB	S	SK	SK	Naik
26	Juli 2018	SK	SK	SK	SB	S	SK	SK	Naik
27	Agustus 2018	SK	SK	SK	SB	S	SK	SK	Turun
28	Oktober 2018	SK	SK	SK	SB	S	SK	SK	Naik
29	Januari 2019	SK	SK	SK	SB	S	SK	SK	Naik
30	Maret 2019	SK	SK	SK	SB	S	SK	SK	Naik
31	April 2019	SK	SK	SK	SB	S	SK	SK	Naik
32	Mei 2019	SK	SK	SK	SB	S	SK	SK	Naik
33	Juli 2019	SK	SK	SK	SB	S	SK	SK	Naik
34	Agustus 2019	SK	SK	SK	SB	S	SK	SK	Turun
35	Oktober 2019	SK	SK	SK	SB	S	SK	SK	Naik
36	Januari 2020	SK	SK	SK	SB	S	SK	SK	Naik
37	Juni 2020	SK	SK	SK	SB	S	SK	SK	Naik
38	Juli 2020	SK	SK	SK	SB	S	SK	SK	Naik

Selanjutnya menghitung:

- Jumlah kasus
- Jumlah kasus keputusan “Naik”
- Jumlah kasus keputusan “Turun”
- Nilai entropy dan gain

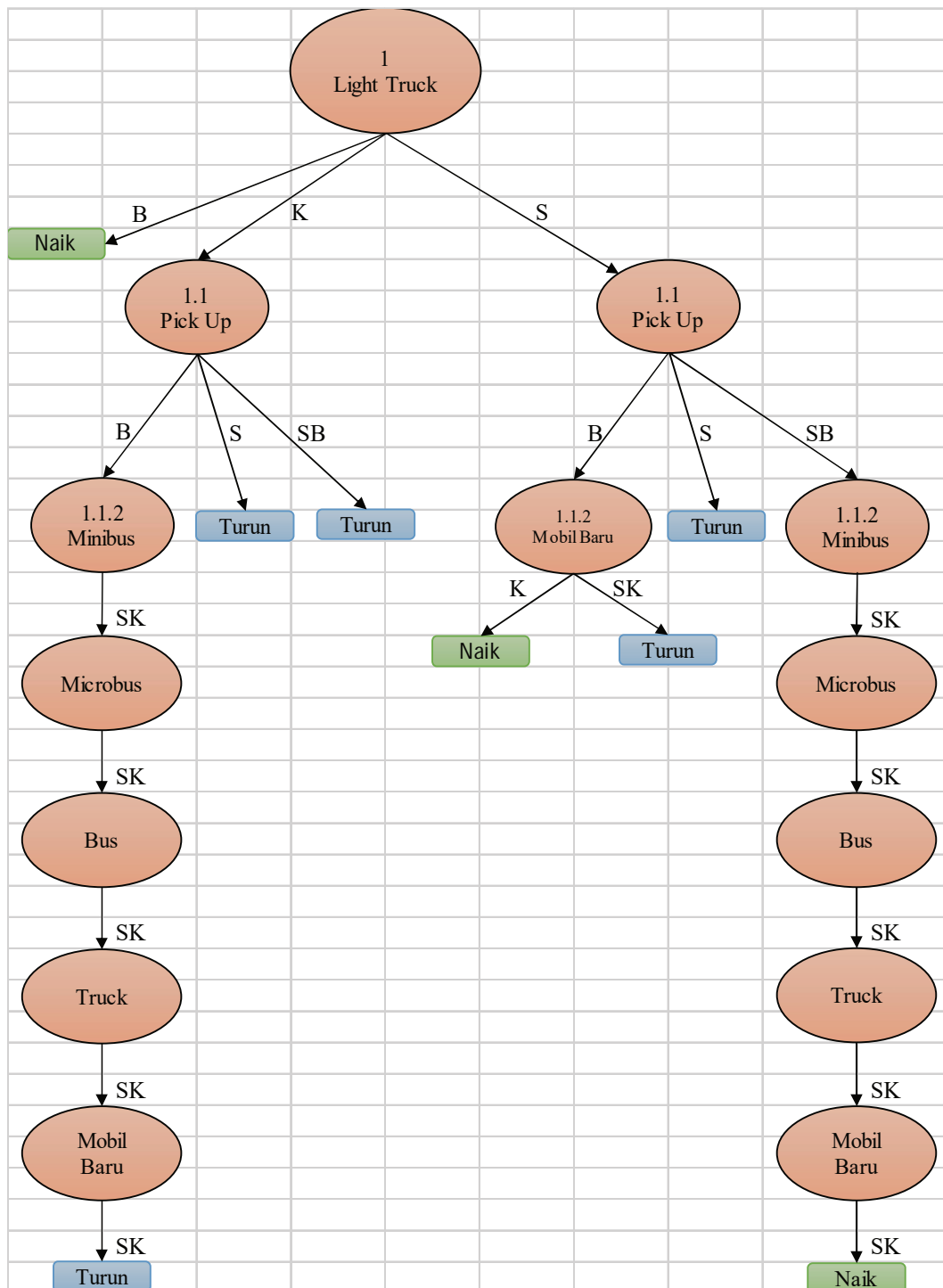
Hasil perhitungan dapat dilihat pada Tabel 3.13 di bawah ini:

Tabel 3.13 Perhitungan Node 1.1.2 Light Truck:S dan Pick Up:SB

Node 1.1.2 Light Truck:S dan Pick Up:SB		Jml Kasus	Naik	Turun	Entropy	Gain
		S	S <sub>1</sub>	S <sub>2</sub>		
Total		38	25	13	0,92682	
Minibus	SK	38	25	13	0,92682	0
Microbus	SK	38	25	13	0,92682	0
Bus	SK	38	25	13	0,92682	0
Truck	SK	38	25	13	0,92682	0
Mobil Baru	SK	38	25	13	0,92682	0

Dari hasil perhitungan Tabel 3.13 terlihat kelima nilai atribut sudah mengklasifikasikan pada satu kasus yaitu keputusannya “Naik” maka tidak perlu dilakukan perhitungan lagi.

Dari hasil proses ke-3 di atas dapat digambarkan pohon keputusan terakhir seperti pada Gambar 3.3 sebagai berikut:



Gambar 3.3 Pohon Keputusan Node 1.1.2

### **3.5. Fase Evaluasi (*Evaluation*)**

Pada tahap evaluasi yang dilakukan dalam penelitian ini adalah mengukur tingkat akurasi dari algoritma c4.5 dalam memodelkan data untuk prediksi pendapatan. Pengukuran data dilakukan dengan confusion matrix, dimana proses perhitungannya menggunakan alat bantu RapidMiner versi 5.3, yang selanjutnya akan dijabarkan pada Bab IV.

### **3.6. Fase Penyebaran (*Deployment*)**

Pada tahap ini, pengetahuan atau informasi yang telah diperoleh akan dipresentasikan dalam bentuk laporan. Hasil dari penelitian ini berupa prediksi naik atau turunnya pendapatan retribusi Uji KIR pada Dinas Perhubungan Kabupaten Karanganyar yang diharapkan dapat digunakan oleh instansi sebagai bahan pertimbangan dalam menentukan target pendapatan pada tahun 2021 yang akan datang.

Hasil dari proses perhitungan algoritma C4.5 secara manual menunjukkan bahwa atribut Light Truck terpilih sebagai akar pohon dan prediksi yang dihasilkan adalah naik.