

IMPLEMENTASI METODE DECISION TREE C4.5 DALAM PEMBERIAN SUBSIDI LISTRIK KEPADA MASYARAKAT

Afian Nugroho¹, Aji Supriyanto²,

^{1,2}Program Studi Teknik Informatika, Fakultas Ilmu Teknologi Informasi, Universitas Stikubank Semarang
e-mail: ¹afianm@gmail.com, ²ajisup@edu.unisbank.ac.id

Abstrak

Subsidi listrik diberikan dengan tujuan agar ketersediaan listrik dapat terpenuhi bagi pelanggan yang kurang mampu dan masyarakat yang belum terjangkau pelayanan PT. PLN dapat ikut menikmati energi listrik. Subsidi listrik ditengarai lebih banyak dinikmati oleh rumah tangga mampu dari sisi ekonomi. Ini disebabkan kebijakan subsidi listrik saat ini adalah subsidi harga, sehingga semakin besar jumlah konsumsi listriknya semakin besar juga jumlah subsidi listrik yang dinikmati. Penelitian ini membangun aplikasi yang dapat digunakan untuk melakukan klasifikasi masyarakat berdasarkan tingkat kemampuan ekonomi menggunakan algoritma decision tree C4.5 dengan tahapan data mining. Tujuannya dapat menentukan klasifikasi masyarakat yang berhak menerima subsidi listrik atau tidak berhak. Total data yang digunakan pada penelitian adalah sebanyak 220 data. Proses data dibagi menjadi 2 bagian yaitu data latih sebanyak 200 data dan data uji sebanyak 20 data dan metode analisis data. Hasil dari pengujian didapatkan dalam melakukan klasifikasi subsidi listrik dengan 200 data latih dan 20 data uji mendapat hasil akurasi sebesar 1 atau 100 %.

Kata Kunci: Subsidi Listrik, Klasifikasi, Decision Tree, C4.5

1. PENDAHULUAN

Listrik merupakan salah satu kebutuhan dasar masyarakat dan merupakan salah satu kebutuhan hajat hidup orang banyak, sehingga perlu diatur dan disediakan oleh negara sesuai amanah undang-undang 1945 pasal 33. Subsidi diberikan dengan tujuan agar ketersediaan listrik dapat terpenuhi, serta membantu pelanggan yang kurang mampu dan masyarakat yang belum terjangkau pelayanan PT. PLN sehingga dapat ikut menikmati energi listrik. Namun pada kenyataannya anggaran subsidi listrik selama ini banyak yang salah sasaran, sebagian besar dana subsidi jatuh ke orang-orang yang mampu secara ekonomi[1]. Permasalahan lain adalah penentuan daya listrik juga dialami di sisi PLN dimana harus menentukan pelanggan yang tepat untuk mendapatkan subsidi listrik yang disediakan pada daya 450 VA dan 900VA[2]. Berdasarkan hasil evaluasi BKF dengan German *International Cooperation* (GIZ) terhadap subsidi listrik yang diberikan kepada kelompok pelanggan R1-450 VA dan R1-900 VA yang berlaku saat ini menunjukkan subsidi listrik tidak tepat sasaran, karena 5,9 juta pelanggan R1-450 VA dan 14,4 juta pelanggan R1-900 VA adalah kelompok rumah tangga yang telah mampu karena termasuk dalam pengeluaran per kapita lebih dari Rp.1 juta per bulan. Hal lain karena disebabkan kebijakan subsidi listrik saat ini adalah subsidi harga, sehingga semakin besar jumlah konsumsi listriknya semakin besar juga jumlah subsidi listrik yang dinikmati termasuk orang mampu[3].

Guna mengurangi risiko salah sasaran tersebut, perlu dilakukan klasifikasi dalam pemberian subsidi listrik berdasarkan kriteria-kriteria tertentu yang telah menjadi standar di PT. PLN menggunakan *data mining* dengan metode *decision tree*[2]. Metode *decision tree* dengan algoritma ID3 dan C4.5 dipilih untuk kalkulasi pada saat proses *training* data. Hasil pengujian kedua algoritma tersebut didapatkan akurasi sebesar 62 % untuk algoritma ID3, sedangkan hasil algoritma C4.5 sebesar 88 %[2]. Oleh karena itu dalam penelitian ini dipilih algoritma C4.5 agar mendapatkan hasil akurasi yang lebih baik.

Klasifikasi adalah proses untuk menemukan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data dengan tujuan untuk memperkirakan kelas yang tidak diketahui dari suatu objek[4]. Pohon keputusan (*decision tree*) memiliki konsep mengubah data menjadi pohon keputusan dan aturan-aturan keputusan. Proses dalam pohon keputusan pertama data diubah menjadi ke bentuk tabel, kemudian dari bentuk tabel diubah kembali menjadi bentuk pohon keputusan dan mengubah pohon keputusan menjadi aturan. Manfaat dari penggunaan pohon keputusan ini adalah kemampuannya untuk dapat mengetahui dan mengambil keputusan yang sangat kompleks sekalipun menjadi lebih mudah dan sederhana sehingga dalam pengambilan keputusan lebih mudah dalam permasalahan[5]. *Data Mining* merupakan proses ataupun kegiatan untuk mengumpulkan data yang berukuran besar kemudian mengekstraksi data tersebut menjadi informasi – informasi yang nantinya dapat digunakan[6]. Tahapan *data mining* terdiri dari *data cleaning*, *data integration*, *data selection*, *data transformation*, *mining process*, *pattern evaluation*, dan *knowledge presentation*[7].

2. METODE PENELITIAN

2.1. Data Penelitian

Data penelitian ini merupakan data sekunder yang diambil pada PLN Unit Pelaksana Pengatur Distribusi (UP2D) kota Semarang yang memiliki data pemberian subsidi listrik masyarakat kota Semarang. Data diambil melalui wawancara pada bagian pelayanan dan teknik PLN UP2D mengenai proses pemberian subsidi listrik kepada masyarakat. Data terpilih yang layak diolah sejumlah 200 pelanggan.

2.2. Metode yang digunakan

Metode analisis yang digunakan dalam penelitian ini menggunakan proses *data mining* dengan metode *decision tree* algoritma C4.5. Tahapan-tahapan yang dilakukan:

- a. Pembersihan data (*data cleaning*). Pembersihan data merupakan proses proses menghilangkan *noise* dan data yang tidak konsisten atau data tidak relevan. Data pemberian subsidi listrik kepada masyarakat tahun 2020 telah konsisten dan relevan satu sama lain.
- b. Integrasi data (*data integration*). Integrasi data merupakan penggabungan data dari berbagai *database* ke dalam suatu *database* baru. Data pemberian subsidi listrik kepada masyarakat tahun 2020 disajikan dengan Microsoft Excel yang berjumlah total 220 data dalam format *xlsx*.
- c. Seleksi data (*data selection*). Data yang ada pada *database* sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari *database*. Atribut yang diseleksi pada tahap ini adalah atribut pekerjaan, atribut penghasilan, atribut aset kendaraan, atribut jumlah tanggungan keluarga, atribut kepemilikan rumah, atribut jenis lantai dan atribut sumber air minum.
- d. Transformasi data (*data transformasi*). Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam *data mining*. Dalam hal ini, data pemberian subsidi listrik kepada masyarakat tahun 2020 dapat langsung diolah dengan menerapkan metode *decision tree* C4.5. Proses transformasi data dibagi menjadi:
 - 1) Atribut penghasilan per bulan dibagi menjadi 3 kriteria yaitu:
 - a) Penghasilan per bulan < 1.500.000 ditransformasikan menjadi “Rendah”.
 - b) Penghasilan per bulan antara 1.500.000 - 2.500.000 ditransformasikan menjadi “Sedang”.
 - c) Penghasilan per bulan > 2.500.000 ditransformasikan menjadi “Tinggi”.
 - 2) Atribut tanggungan keluarga dibagi menjadi 3 kriteria yaitu :
 - a) Tanggungan keluarga = 1 ditransformasikan menjadi “Sedikit”.
 - b) Tanggungan keluarga = 2 ditransformasikan menjadi “Sedang”.
 - c) Tanggungan keluarga ≥ 3 ditransformasikan menjadi “Banyak”.
- e. Proses *mining*. Proses utama saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data. Tahap ini menerapkan metode *decision tree* C4.5 dalam menentukan klasifikasi pemberian subsidi listrik kepada masyarakat tahun 2020. Tahapan implementasi metode *decision tree* C4.5 dalam pemberian subsidi listrik kepada masyarakat yaitu :
 - 1) Atribut pemberian subsidi adalah atribut pekerjaan, atribut penghasilan, atribut aset kendaraan, atribut jumlah tanggungan keluarga, atribut kepemilikan rumah, atribut jenis lantai dan atribut sumber air minum. Setiap atribut memiliki nilai sedangkan kelasnya ada pada kolom hasil yaitu kelas subsidi dan kelas nonsubsidi.
 - 2) Hitung entropi dengan rumus :

$$\text{Entropi (S)} = \sum_{j=1}^k -p_j \log_2 p_j \quad (1)$$
 - 3) Setelah mendapatkan entropi dari keseluruhan kasus, lakukan analisis setiap atribut dan nilai-nilainya dan hitung entropinya.
 - 4) Hitung gain untuk setiap atribut dengan rumus :

$$\text{Gain (A)} = \text{Entropi (S)} - \sum_{i=1}^k \frac{|S_i|}{|S|} \times \text{Entropi (S}_i) \quad (2)$$
 - 5) Gain terbesar akan menjadi node akar (*root node*).
 - 6) Lakukan lagi langkah-langkah seperti sebelumnya sampai didapatkan pohon keputusan.
- f. Evaluasi pola (*pattern evaluation*). Proses mengidentifikasi pola-pola menarik ke dalam *knowledge based* yang ditemukan. Tahap ini menentukan pola dari implementasi metode *decision tree* C4.5 dalam pemberian subsidi listrik kepada masyarakat.
- g. Presentasi pengetahuan (*knowledge presentation*). Presentasi pengetahuan akan menampilkan visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna. Tahap ini menampilkan visualisasi pohon keputusan dari implementasi metode *decision tree* C4.5 dalam pemberian subsidi listrik kepada masyarakat.

3. HASIL DAN PEMBAHASAN

3.1. Hasil

Pada penelitian ini melakukan klasifikasi masyarakat ke dalam masyarakat yang berhak menerima subsidi listrik atau masyarakat yang tidak berhak menerima subsidi listrik menggunakan metode *decision tree* dengan algoritma *C4.5*. Proses klasifikasi menggunakan *tool* Rstudio yang pada penerapannya pada *decision tree* membutuhkan *packages* yaitu *fselector*, *rpart*, *caret*, *dplyr*, *rpart.plot*, *readxl*, *data.tree* dan *caTools* yang dipanggil seperti pada Gambar 1.

```
library(fselector)
library(rpart)
library(caret)
library(dplyr)
library(rpart.plot)
library(readxl)
library(data.tree)
library(caTools) ...
```

Gambar 1. Script Pemanggilan Packages

Proses selanjutnya adalah membaca data excel subsidi listrik dengan tipe *.xlsx* seperti Gambar 2. Data subsidi listrik masyarakat dengan tipe *xlsx* selanjutnya akan dipanggil dengan nama *df*. Hal ini dilakukan untuk mempermudah pemanggilan data subsidi listrik (*df*) dalam proses klasifikasi menggunakan *decision tree*.

	pekerjaan	penghasilan	aset	tanggungan	rumah	lantai	air	hasil
1	SERABUTAN	Rendah	SEPEDA	BANYAK	KELUARGA	TANAH	SUMUR	SUBSIDI
2	PEGAWAI	Tinggi	MOBIL	BANYAK	MILIK SENDIRI	KERAMIK	PDAM	NONSUBSIDI
3	WIRUSAHA	Tinggi	MOBIL	BANYAK	SEWA	PLESTER	PDAM	NONSUBSIDI
4	BURUH	Sedang	MOTOR	BANYAK	SEWA	PLESTER	PDAM	SUBSIDI
5	SERABUTAN	Rendah	TIDAK MEMILIKI	BANYAK	KELUARGA	TANAH	SUMUR	SUBSIDI
6	WIRUSAHA	Sedang	MOTOR	BANYAK	MILIK SENDIRI	PLESTER	PDAM	SUBSIDI
7	PEGAWAI	Tinggi	MOTOR	BANYAK	KELUARGA	MARMER	PDAM	NONSUBSIDI
8	BURUH	Tinggi	MOBIL	SEDIKIT	SEWA	KERAMIK	PDAM	NONSUBSIDI
9	WIRUSAHA	Sedang	MOBIL	SEDANG	MILIK SENDIRI	KERAMIK	PDAM	NONSUBSIDI
10	WIRUSAHA	Rendah	MOTOR	BANYAK	KELUARGA	PLESTER	PDAM	SUBSIDI
11	WIRUSAHA	Sedang	MOBIL	SEDANG	MILIK SENDIRI	KERAMIK	PDAM	NONSUBSIDI
12	BURUH	Rendah	MOTOR	SEDIKIT	MILIK SENDIRI	PLESTER	PDAM	SUBSIDI

```
10 df <- read_excel("D:/afi/databaru.xlsx")
11 str(df)
12 <
28:1 (Top Level)
R Script

Console Terminal
> library(rpart.plot)
> library(readxl)
> library(data.tree)
> library(caTools)
> df <- read_excel("D:/afi/databaru.xlsx")
> str(df)
tibble [220 x 8] (s3: tbl_df/tbl/data.frame)
 $ pekerjaan : chr [1:220] "SERABUTAN" "PEGAWAI" "WIRUSAHA" "BURUH" ...
 $ penghasilan: chr [1:220] "RENDAH" "TINGGI" "TINGGI" "SEDANG" ...
 $ aset       : chr [1:220] "SEPEDA" "MOBIL" "MOBIL" "MOTOR" ...
 $ tanggungan: chr [1:220] "BANYAK" "BANYAK" "BANYAK" "BANYAK" ...
 $ rumah     : chr [1:220] "KELUARGA" "MILIK SENDIRI" "SEWA" "SEWA" ...
 $ lantai    : chr [1:220] "TANAH" "KERAMIK" "PLESTER" "PLESTER" ...
 $ air       : chr [1:220] "SUMUR" "PDAM" "PDAM" "PDAM" ...
 $ hasil     : chr [1:220] "SUBSIDI" "NONSUBSIDI" "NONSUBSIDI" "SUBSIDI" ...
```

Gambar 2. Hasil Membaca Data Subsidi Listrik

Data subsidi listrik tersebut selanjutnya dibagi menjadi data latih dan data uji, *source code* seperti Gambar 3. Data latih digunakan untuk proses pembentukan pohon keputusan, sedangkan data uji digunakan untuk menguji akurasi dari pohon keputusan yang telah terbentuk.

```
df <- select(df,pekerjaan, penghasilan, aset, tanggungan, rumah, lantai, air, hasil)
df <- mutate(df, pekerjaan = factor(pekerjaan),penghasilan = factor(penghasilan),aset = fact

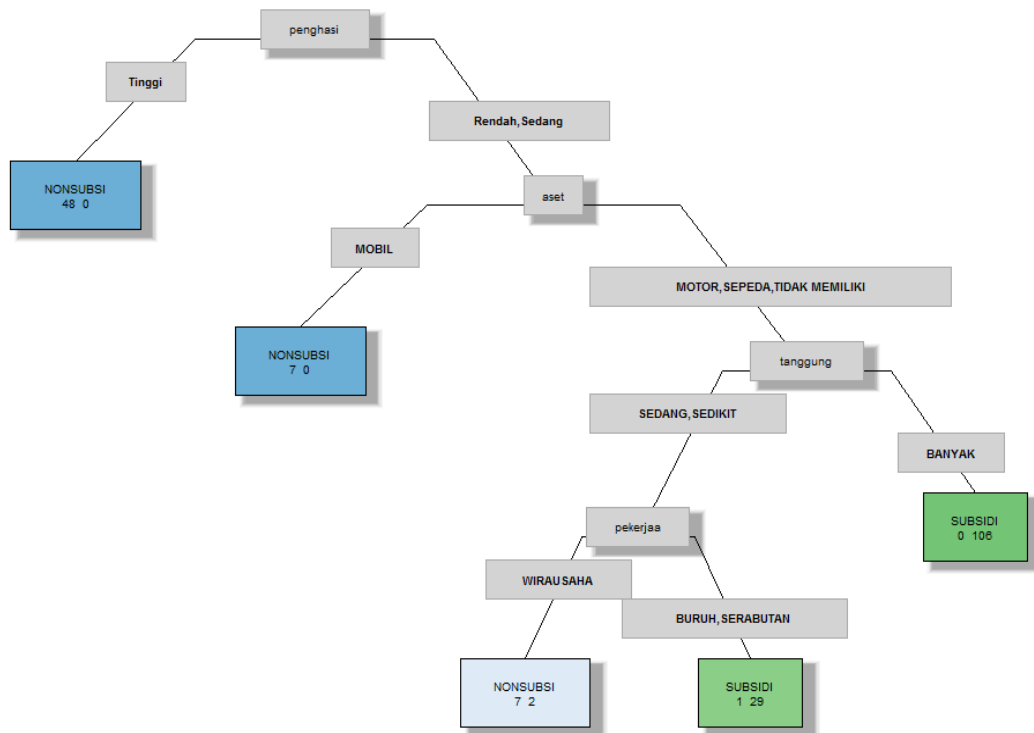
train <- df[1:200,]
test <- df[201:220,]
```

Gambar 3. Script Pembagian Data

Pada 220 data subsidi listrik, data latih yang dipakai adalah 200 data, sedangkan data uji adalah 20 data. Proses pembagian data latih dan data uji dilakukan dengan *trial and error* menggunakan bahasa R dan didapatkan hasil pembentukan pohon terbaik dengan menggunakan data latih sebanyak 200 data dari total data. Proses pembagian data latih dan data uji dilakukan secara *sequential* atau berurut dimana data latih terletak pada baris 1 sampai dengan baris 200 dari data subsidi listrik. Berikutnya data uji terletak pada baris 201 sampai dengan baris 220 dari data subsidi listrik.

Proses selanjutnya adalah pembentukan *tree* dengan *class* hasil yaitu subsidi dan non subsidi. Data latih dan pohon keputusan yang berhasil terbentuk dari data latih akan diperlihatkan seperti Gambar 4. Pohon keputusan yang terbentuk akan menampilkan hasil keputusan subsidi atau non subsidi beserta jumlah data yang mewakili hasil keputusan terbentuknya keputusan subsidi dan non subsidi. Sebagai contoh penghasilan

tinggi akan menghasilkan keputusan non subsidi dengan jumlah data dengan hasil non subsidi sebanyak 65 data dan keputusan subsidi dengan jumlah data dengan hasil subsidi sebanyak 0 data.



Gambar 4. Pohon Keputusan Data Latih

Pohon keputusan seperti Gambar 4 menghasilkan beberapa keputusan sebagai berikut:

- Jika penghuni mempunyai penghasilan tinggi maka akan diklasifikasikan ke dalam golongan non subsidi.
- Jika penghuni mempunyai penghasilan sedang atau rendah dan mempunyai aset berupa mobil maka akan diklasifikasikan ke dalam golongan non subsidi.
- Jika penghuni mempunyai penghasilan sedang atau rendah, memiliki aset berupa motor atau sepeda atau tidak memiliki aset dan mempunyai tanggungan yang banyak maka akan diklasifikasikan ke dalam golongan subsidi.
- Jika penghuni mempunyai penghasilan sedang atau rendah, memiliki aset berupa motor atau sepeda atau tidak memiliki aset dan pekerjaannya adalah wirasaha maka akan diklasifikasikan ke dalam golongan non subsidi.
- Jika penghuni mempunyai penghasilan sedang atau rendah, memiliki aset berupa motor atau sepeda atau tidak memiliki aset dan pekerjaannya adalah buruh atau serabutan maka akan diklasifikasikan ke dalam golongan subsidi.

Script untuk membentuk pohon keputusan terlihat pada Gambar 5.

```
tree <- rpart(hasil ~ ., data= train)
prp(tree, extra = 1, type = 5, box.palette="auto",shadow.col = "darkgray",
    facLen = 0,branch = .5,split.box.col = "lightgray",split.border.col = "darkgray",
    compress = FALSE,leaf.round=0)
```

Gambar 5. Script Pembentukan Pohon Keputusan

Proses pengujian dari data uji dilakukan dengan menggunakan *confusion matrix* seperti Gambar 6.

```
tree.hasil.predicted<-predict(tree,test,type='class')
confusionMatrix(tree.hasil.predicted, test$hasil)
```

Gambar 6. Script Pengujian *Confusion Matrix*

Hasil akurasi metode *decision tree* dalam menentukan klasifikasi masyarakat yang berhak menerima subsidi listrik atau masyarakat yang tidak berhak menerima subsidi listrik dihitung dengan *confusion matrix* menghasilkan akurasi 1 atau 100 % yang diperlihatkan seperti Gambar 7.

		Reference	
Prediction		NONSUBSIDI	SUBSIDI
NONSUBSIDI		8	0
SUBSIDI		0	12

Accuracy : 1
 95% CI : (0.8316, 1)
 No Information Rate : 0.6
 P-Value [Acc > NIR] : 3.656e-05
 Kappa : 1
 McNemar's Test P-Value : NA
 Sensitivity : 1.0
 Specificity : 1.0
 Pos Pred Value : 1.0
 Neg Pred Value : 1.0
 Prevalence : 0.4
 Detection Rate : 0.4
 Detection Prevalence : 0.4
 Balanced Accuracy : 1.0
 'Positive' Class : NONSUBSIDI

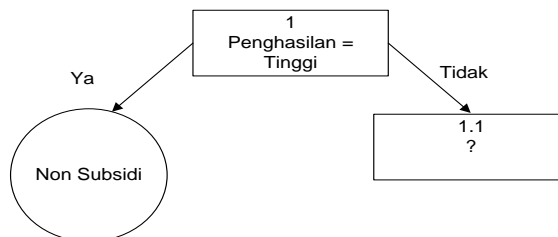
Gambar 7. Hasil Akurasi Klasifikasi Data Uji

Berdasarkan Gambar 7, hasil pengujian klasifikasi data subsidi listrik menggunakan algoritma *decision tree* dengan dengan *confusion matrix* menghasilkan :

- a. Data non subsidi yang terdeteksi menjadi hasil non subsidi sebanyak 8 data.
- b. Data non subsidi yang terdeteksi menjadi hasil subsidi sebanyak 0 data.
- c. Data subsidi yang terdeteksi menjadi hasil non subsidi sebanyak 0 data
- d. Data subsidi yang terdeteksi menjadi hasil subsidi sebanyak 12 data

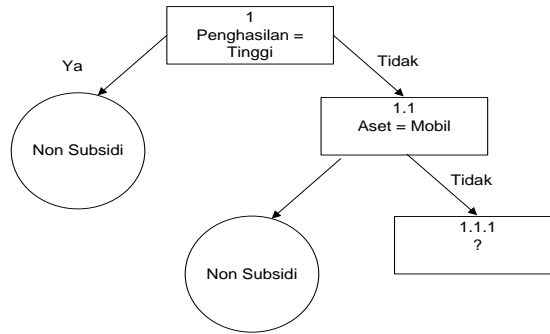
3.2. Pembahasan

Proses klasifikasi pemberian subsidi listrik kepada masyarakat tahun 2020 menggunakan metode *decision tree* C4.5 diklasifikasikan menjadi 2 klasifikasi yaitu subsidi dan non subsidi. Klasifikasi berdasarkan pekerjaan, penghasilan, aset kendaraan, jumlah tanggungan keluarga, kepemilikan rumah, jenis lantai dan sumber air minum. Data latih memiliki 200 kasus yang terdiri dari 137 data subsidi dan 63 data non subsidi. Nilai *gain* terbesar pada *node* pertama adalah *gain* (penghasilan), maka penghasilan menjadi *node* akar (*root node*). Penghasilan tinggi memiliki 48 kasus dan semuanya memiliki jawaban non subsidi. Penghasilan tinggi akan menjadi daun atau *leaf*. Pohon keputusan yang terbentuk pada *node* 1 diperlihatkan seperti Gambar 8.



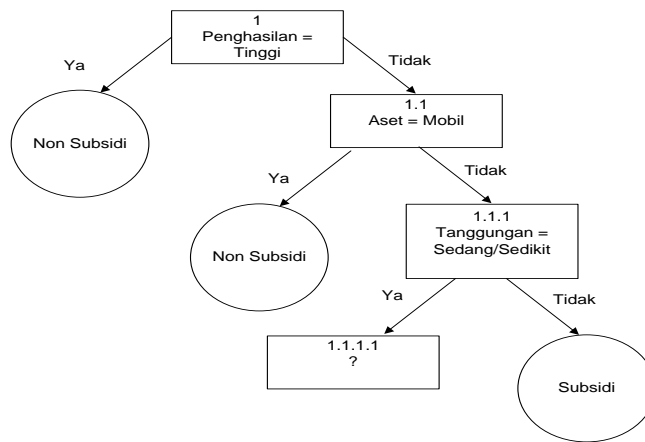
Gambar 8. Pohon Keputusan dengan *Node* 1

Nilai *gain* terbesar berikutnya pada *node* 1.1 adalah *gain* (aset), maka aset menjadi *node* berikutnya. Aset mobil memiliki 7 kasus dan semuanya memiliki jawaban non subsidi. Aset mobil selanjutnya menjadi daun atau *leaf*. Pohon keputusan yang terbentuk pada *node* 1.1 diperlihatkan seperti Gambar 9.



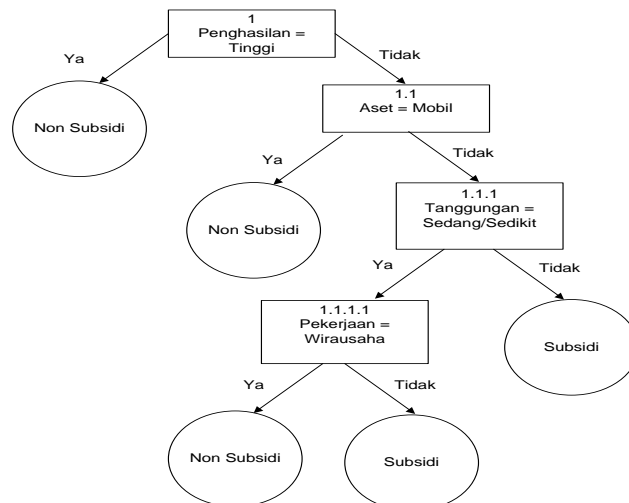
Gambar 9. Pohon Keputusan Node 1.1

Nilai *gain* terbesar berikutnya pada node 1.1.1 adalah *gain* (tanggungan), maka tanggungan menjadi node berikutnya. Tanggungan banyak memiliki 106 kasus dan semuanya memiliki jawaban subsidi. Tanggungan banyak selanjutnya menjadi daun atau *leaf*. Pohon keputusan yang terbentuk pada node 1.1.1 diperlihatkan seperti Gambar 10.



Gambar 10. Pohon Keputusan Node 1.1.1

Nilai *gain* terbesar berikutnya pada node 1.1.1.1 adalah *gain* (pekerjaan), maka tanggungan menjadi node berikutnya. Pekerjaan akan menjadi daun atau *leaf*. Pohon keputusan yang terbentuk pada node 1.1.1.1 atau pohon keputusan akhir diperlihatkan seperti Gambar 11.



Gambar 11. Pohon Keputusan Akhir

Gambar 11 merupakan pohon keputusan akhir yang terbentuk dari implementasi metode *decision tree* C4.5 dalam pemberian subsidi listrik kepada masyarakat. Pohon keputusan seperti Gambar 11 selanjutnya dapat menghasilkan 5 keputusan seperti Tabel 1.

Tabel 1. Tabel Keputusan Subsidi Listrik

No	Penghasilan	Aset	Tanggungsan	Pekerjaan	Hasil
1.	Tinggi	-	-	-	Nonsubsidi
2.	Rendah/Sedang	Mobil	-	-	Nonsubsidi
3.	Rendah/Sedang	Motor/Sepeda/ Tidak Memiliki	Banyak	-	Subsidi
4.	Rendah/Sedang	Motor/Sepeda/ Tidak Memiliki	Sedang/Sedikit	Wirausaha	Nonsubsidi
5	Rendah/Sedang	Motor/Sepeda/ Tidak Memiliki	Sedang/Sedikit	Buruh/ Serabutan	Subsidi

Setelah didapatkan tabel keputusan seperti Tabel 1 selanjutnya dilakukan pengujian, dan dari data uji yang terbentuk seperti terlihat pada Tabel 2.

Tabel 2. Hasil Pengujian Subsidi Listrik

No	Pekerjaan	Penghasilan	Aset	Tanggungsan	Hasil	Klasifikasi
1.	Wirausaha	Tinggi	Mobil	Banyak	Nonsubsidi	Nonsubsidi
2.	Wirausaha	Sedang	Motor	Banyak	Subsidi	Subsidi
3.	Wirausaha	Sedang	Motor	Banyak	Subsidi	Subsidi
4.	Buruh	Sedang	Tidak Memiliki	Banyak	Subsidi	Subsidi
5.	Buruh	Sedang	Motor	Banyak	Subsidi	Subsidi
6.	Wirausaha	Sedang	Motor	Banyak	Subsidi	Subsidi
7.	Wirausaha	Sedang	Motor	Sedikit	Nonsubsidi	Nonsubsidi
8.	Buruh	Sedang	Tidak Memiliki	Sedang	Subsidi	Subsidi
9.	Serabutan	Rendah	Tidak Memiliki	Banyak	Subsidi	Subsidi
10.	Wirausaha	Sedang	Motor	Banyak	Subsidi	Subsidi
11.	Pegawai	Tinggi	Mobil	Banyak	Nonsubsidi	Nonsubsidi
12.	Buruh	Tinggi	Mobil	Sedikit	Nonsubsidi	Nonsubsidi
13.	Buruh	Tinggi	Mobil	Banyak	Nonsubsidi	Nonsubsidi
14.	Wirausaha	Sedang	Mobil	Sedikit	Nonsubsidi	Nonsubsidi
15.	Buruh	Sedang	Motor	Sedang	Subsidi	Subsidi
16.	Wirausaha	Sedang	Motor	Banyak	Subsidi	Subsidi
17.	Pegawai	Tinggi	Motor	Banyak	Nonsubsidi	Nonsubsidi
18.	Buruh	Sedang	Motor	Sedang	Subsidi	Subsidi
19.	Buruh	Sedang	Motor	Banyak	Subsidi	Subsidi
20.	Wirausaha	Sedang	Mobil	Sedang	Nonsubsidi	Nonsubsidi

Hasil klasifikasi pada Tabel 2 menghasilkan klasifikasi sebagai berikut:

- Klasifikasi nonsubsidi dari hasil nonsubsidi menghasilkan klasifikasi sebesar 8 klasifikasi yang diberi warna biru atau disebut TP (*true negatif*).
- Klasifikasi subsidi dari hasil subsidi menghasilkan klasifikasi sebesar 12 klasifikasi yang diberi warna hijau pada tabel atau disebut TP (*true positif*).
- Klasifikasi nonsubsidi dari hasil subsidi menghasilkan klasifikasi sebesar 0 klasifikasi atau disebut FN (*false negatif*).
- Klasifikasi subsidi dari hasil nonsubsidi menghasilkan klasifikasi sebesar 0 klasifikasi atau disebut FP (*false positif*).

Hasil akurasi metode *decision tree* dalam menentukan klasifikasi masyarakat yang berhak menerima subsidi listrik atau masyarakat yang tidak berhak menerima subsidi listrik dihitung dengan menggunakan *confusion matrix* yang rumusnya :

$$Akurasi = \frac{TN+TP}{TN+TP+FN+FP} \tag{3}$$

Sehingga menghasilkan akurasi sebesar :

$$Akurasi = \frac{8+12}{8+12+0+0}$$

$$Akurasi = \frac{20}{20} = 1$$

Akurasi algoritma *decision tree* dalam melakukan klasifikasi masyarakat yang berhak menerima subsidi listrik atau masyarakat yang tidak berhak menerima subsidi listrik adalah sebesar 1 atau sebesar 100%.

4. KESIMPULAN

Penelitian ini menghasilkan klasifikasi masyarakat yang berhak menerima subsidi dan non subsidi listrik dengan nilai akurasi 1 atau 100 % dengan menggunakan metode *decision tree C4.5*. Hasil tersebut perlu dilakukan pengujian kembali dengan data yang lebih besar, beragam dan menggunakan data primer untuk menunjukkan hasil yang lebih nyata. Selain itu pengujian juga perlu dilakukan dengan metode lain untuk menunjukkan perbandingan yang lebih baik.

DAFTAR PUSTAKA

- [1] Widarma, A., & Kumala, H., 2018, Sistem Pendukung Keputusan Dalam Menentukan Pengguna Listrik Subsidi Dan Nonsubsidi Menggunakan Metode Fuzzy Mamdani (Studi Kasus: PT. PLN Tanjung Balai). *JurTI (Jurnal Teknologi Informasi)*, 2(2), 165-171.
- [2]. Tanjung, Y. P., Sentinuwo, S. R., & Jacobus, A., 2016, Penentuan Daya Listrik Rumah Tangga Menggunakan Metode Decision Tree. *Jurnal Teknik Informatika*, 9(1).
- [3]. Badan Pusat Statistik., 2015, *Indonesia - Survei Sosial Ekonomi Nasional 2014 (Gabungan)*, Jakarta
- [4] Han, J., & Kamber, M., 2012, *Data Mining Concepts and Techniques 3rd Edition*, Waltham, United Kingdom
- [5] Kusrini & Luthfi, E.T., 2009, *Algoritma Data Mining*, Andi, Yogyakarta
- [6] Saleh, A., 2015, Implementasi metode klasifikasi naive bayes dalam memprediksi besarnya penggunaan listrik rumah tangga. *Creative Information Technology Journal*, 2(3), 207-217.
- [7]. Ridwan, M., Suyono, H., & Sarosa, M., 2013, Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier. *jurnal EECCIS*, 7(1), 59-64.