

# Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga

**Alfa Saleh**

Universitas Potensi Utama

E-mail: [alfasoleh1@gmail.com](mailto:alfasoleh1@gmail.com)

## **Abstrak**

*Peranan listrik sangat penting bagi setiap lapisan masyarakat bahkan listrik juga sangat dibutuhkan sebagai sarana produksi dan untuk kehidupan sehari-hari, begitu pentingnya peranan listrik tentu saja berdampak pada permintaan listrik yang semakin besar tapi hal ini kiranya tidak linier dengan persediaan listrik yang belum mampu memenuhi permintaan listrik yang begitu besar tersebut. Untuk mengatasi hal ini perlu adanya campur tangan pemerintah dan masyarakat dalam menggunakan listrik dengan bijak sehingga kebutuhan listrik tidak menjadi lebih besar dari persediaan listrik. Oleh karena itu setiap rumah tangga haruslah paham penggunaan listrik yang efektif. Penerapan metode naïve bayes diharapkan mampu untuk memprediksi besarnya penggunaan listrik tiap rumah tangga agar lebih mudah mengatur penggunaan listrik. dari 60 data penggunaan listrik rumah tangga yang diuji dengan metode naïve bayes, maka diperoleh hasil persentase 78,3333% untuk keakuratan prediksi, di mana dari 60 data penggunaan listrik rumah tangga yang diuji terdapat 47 data penggunaan listrik rumah tangga yang berhasil diklasifikasikan dengan benar.*

**Kata Kunci** — Data Mining, Naïve Bayes, Penggunaan Listrik.

## **Abstract**

*The role of electricity is very important for every layer of society and even electricity is also needed as a means of production and to everyday in life, the importance of the role of electricity of course have an impact on the growing electricity demand but it would not be linear with electricity supply not been able to meet demand the electricity is so great. To address this need for government intervention and community to use electricity wisely so that the electricity needs not be greater than the supply of electricity. Therefore, every household should be understood that the effective use of electricity. Implementation of Naïve Bayes method is able to predict the magnitude of expected electricity use per household in order to more easily manage the use of electricity. Based on 60 household electricity usage data tested with Naïve Bayes method, be obtained the percentage 78.3333% for the accuracy of the prediction, in which of the 60 household electricity usage data are tested, there are 47 household electricity usage data successfully classified correctly.*

**Keywords** — Data Mining, Naïve Bayes, Electricity Usage.

---

## 1. PENDAHULUAN

### 1.1. Latar Belakang

Kemajuan teknologi di segala bidang meningkat dengan begitu cepat, kemajuan ini membawa konsekuensi peningkatan kebutuhan akan daya listrik. Listrik merupakan salah bentuk energi yang banyak dibutuhkan, ini dimungkinkan karena energi listrik mudah dalam penyaluran dan dapat dengan mudah dirubah ke bentuk energi [1]. Listrik menjadi kebutuhan pokok bagi masyarakat baik dalam bidang industri maupun rumah tangga. Penggunaan energi listrik di rumah tangga terutama untuk penerangan, alat setrika, hiburan, kipas angin, lemari es dan pendingin ruangan. Penggunaan alat-alat listrik memerlukan arus listrik yang dihasilkan dari sumber energi[2]. Penggunaan listrik yang tidak bijak tentu saja akan berdampak pada tingginya penggunaan listrik, hal ini juga mempengaruhi menipisnya persediaan energi listrik dikarenakan kebutuhan akan energi listrik lebih besar dari persediaan akan energi listrik, untuk itu diharapkan setiap masyarakat memahami upaya dalam menggunakan listrik dengan bijak. Selain itu setiap rumah tangga juga harus tahu besarnya penggunaan listrik di rumah tangga masing-masing. Dalam hal mengetahui penggunaan listrik ini maka diterapkan teknik klasifikasi data mining menggunakan metode naïve bayes dalam memprediksi penggunaan listrik rumah tangga.

Pada penelitian sebelumnya metode *Naive Bayes* juga digunakan dalam memprediksi penyakit Dermatologi yang diabaikan tapi bahkan dapat menyebabkan kematian di mana metode *Naive Bayes* digunakan untuk mengenal pola data untuk mengungkap kemungkinan penyakit dermatologi[3]. Metode *Naive Bayes* juga dinilai berpotensi baik dalam mengklasifikasi dokumen dibandingkan metode pengklasifikasian yang lain dalam hal akurasi dan efisiensi komputasi [4].

### 1.2. Data Mining

*Data Mining* merupakan proses pengekstraksian informasi dari sekumpulan data yang sangat besar melalui penggunaan algoritma dan teknik penarikan dalam bidang statistik, pembelajaran mesin dan sistem manajemen basis data[5]. *Data mining* adalah proses menganalisa data dari perspektif yang berbeda dan menyimpulkannya menjadi informasi-informasi penting yang dapat dipakai untuk meningkatkan keuntungan, memperkecil biaya pengeluaran, atau bahkan keduanya[6]. Definisi lain mengatakan *Data Mining* adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam data berukuran besar[7]. Dari beberapa definisi di atas dapat ditarik kesimpulan bahwa *Data Mining* merupakan proses ataupun kegiatan untuk mengumpulkan data yang berukuran besar kemudian mengekstraksi data tersebut menjadi informasi – informasi yang nantinya dapat digunakan.

#### 4.1.1. Tahap-tahap Data Mining

Sebagai suatu rangkaian proses, *Data Mining* dapat dibagi menjadi beberapa tahap proses. Tahap-tahap tersebut bersifat interaktif, pemakai terlibat langsung atau dengan perantara *knowledge base*.

Tahap-tahap *Data Mining* adalah sebagai berikut[8]:

- a. Pembersihan data (*Data Cleaning*)  
Pembersihan data merupakan proses menghilangkan-kan *noise* dan data yang tidak konsisten atau data tidak relevan.
  - b. Integrasi data (*Data Integration*)  
Integrasi data merupakan penggabungan data dari berbagai *database* ke dalam satu *database* baru.
  - 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*.
-

- d. Transformasi data (Data Transformation)  
Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam *Data Mining*.
- e. Proses *Mining*  
Merupakan suatu proses utama saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data. Beberapa metode yang dapat digunakan berdasarkan pengelompokan *Data Mining*.
- f. Evaluasi pola (*Pattern Evaluation*)  
Untuk mengidentifikasi pola-pola menarik ke dalam *knowledge based* yang ditemukan.
- g. Presentasi pengetahuan (*Knowledge Presentation*)  
Merupakan visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna.

### 1.3. Metode Naive Bayes

Naive Bayes merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Algoritma menggunakan teorema Bayes dan mengasumsikan semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas[9]. Definisi lain mengatakan Naive Bayes merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya [10].

*Naive Bayes* didasarkan pada asumsi penyederhanaan bahwa nilai atribut secara kondisional saling bebas jika diberikan nilai output. Dengan kata lain, diberikan nilai output, probabilitas mengamati secara bersama adalah produk dari probabilitas individu[8]. *Keuntungan penggunaan Naive Bayes adalah bahwa metode ini hanya membutuhkan jumlah data pelatihan (Training Data) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. Naive Bayes sering bekerja jauh lebih baik dalam kebanyakan situasi dunia nyata yang kompleks dari pada yang diharapkan[11].*

#### 4.1.2. Persamaan Metode Naive Bayes

Persamaan dari teorema *Bayes* adalah[10] :

$$P(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)} \quad (1)$$

Di mana :

$X$  : Data dengan *class* yang belum diketahui

$H$  : Hipotesis data merupakan suatu *class* spesifik

$P(H|X)$  : Probabilitas hipotesis  $H$  berdasar kondisi  $X$  (posteriori probabilitas)

$P(H)$  : Probabilitas hipotesis  $H$  (prior probabilitas)

$P(X|H)$  : Probabilitas  $X$  berdasarkan kondisi pada hipotesis  $H$

$P(X)$  : Probabilitas  $X$

Untuk menjelaskan metode *Naive Bayes*, perlu diketahui bahwa proses klasifikasi memerlukan sejumlah petunjuk untuk menentukan kelas apa yang cocok bagi sampel yang dianalisis tersebut. Karena itu, metode *Naive Bayes* di atas disesuaikan sebagai berikut:

$$P(C|F1 \dots Fn) = \frac{P(C)P(F1 \dots Fn|C)}{P(F1 \dots Fn)} \quad (2)$$

Di mana Variabel  $C$  merepresentasikan kelas, sementara variabel  $F_1 \dots F_n$  merepresentasikan karakteristik petunjuk yang dibutuhkan untuk melakukan klasifikasi. Maka rumus tersebut menjelaskan bahwa peluang masuknya sampel karakteristik tertentu dalam kelas  $C$  (*Posterior*) adalah peluang munculnya kelas  $C$  (sebelum masuknya sampel tersebut, seringkali disebut *prior*), dikali dengan peluang kemunculan karakteristik-karakteristik sampel pada kelas  $C$  (disebut juga *likelihood*), dibagi dengan peluang kemunculan karakteristik-karakteristik sampel secara global (disebut juga *evidence*). Karena itu, rumus di atas dapat pula ditulis secara sederhana sebagai berikut:

$$\text{Posterior} = \frac{\text{prior} \times \text{likelihood}}{\text{evidence}} \quad (3)$$

Nilai *Evidence* selalu tetap untuk setiap kelas pada satu sampel. Nilai dari *posterior* tersebut nantinya akan dibandingkan dengan nilai-nilai *posterior* kelas lainnya untuk menentukan ke kelas apa suatu sampel akan diklasifikasikan. Penjabaran lebih lanjut rumus *Bayes* tersebut dilakukan dengan menjabarkan  $(C|F_1, \dots, F_n)$  menggunakan aturan perkalian sebagai berikut:

$$\begin{aligned} P(C|F_1, \dots, F_n) &= P(C)P(F_1, \dots, F_n|C) \\ &= P(C)P(F_1|C)P(F_2, \dots, F_n|C, F_1) \\ &= P(C)P(F_1|C)P(F_2|C, F_1)P(F_3, \dots, F_n|C, F_1, F_2) \\ &= P(C)P(F_1|C)P(F_2|C, F_1)P(F_3|C, F_1, F_2)P(F_4, \dots, F_n|C, F_1, F_2, F_3) \\ &= P(C)P(F_1|C)P(F_2|C, F_1)P(F_3|C, F_1, F_2) \dots P(F_n|C, F_1, F_2, F_3, \dots, F_{n-1}) \end{aligned} \quad (4)$$

Dapat dilihat bahwa hasil penjabaran tersebut menyebabkan semakin banyak dan semakin kompleksnya faktor - faktor syarat yang mempengaruhi nilai probabilitas, yang hampir mustahil untuk dianalisa satu persatu. Akibatnya, perhitungan tersebut menjadi sulit untuk dilakukan. Di sinilah digunakan asumsi independensi yang sangat tinggi (*naif*), bahwa masing-masing petunjuk ( $F_1, F_2 \dots F_n$ ) saling bebas (*independen*) satu sama lain. Dengan asumsi tersebut, maka berlaku suatu kesamaan sebagai berikut:

$$P(F_i|F_j) = \frac{P(F_i \cap F_j)}{P(F_j)} = \frac{P(F_i)P(F_j)}{P(F_j)} = P(F_i) \quad (5)$$

Untuk  $i \neq j$ , sehingga

$$P(F_i|C, F_j) = P(F_i|C) \quad (6)$$

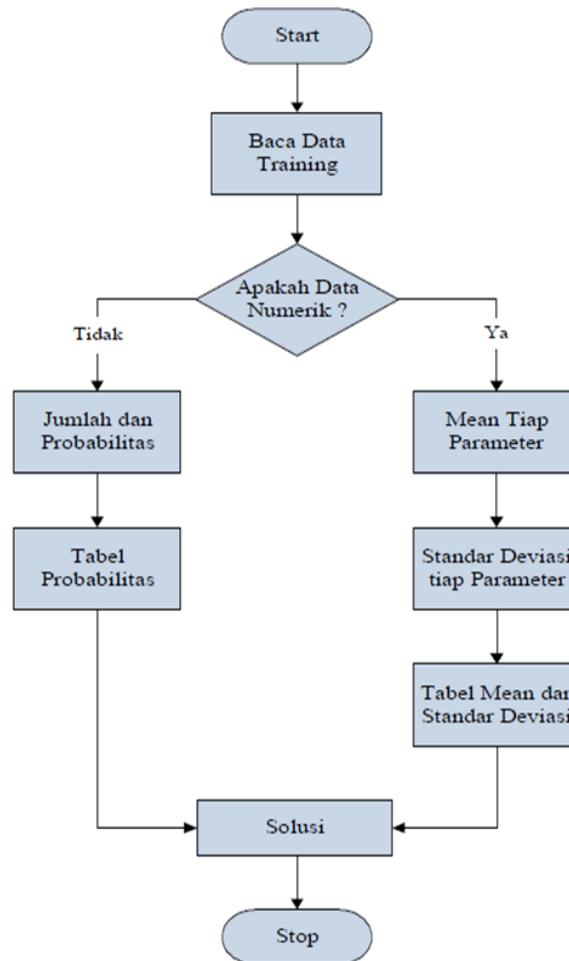
Persamaan di atas merupakan model dari teorema *Naive Bayes* yang selanjutnya akan digunakan dalam proses klasifikasi. Untuk klasifikasi dengan data kontinyu digunakan rumus *Densitas Gauss* :

$$P(X_i = x_i|Y = y_j) = \frac{1}{\sqrt{2\pi\sigma_{ij}}} e^{-\frac{(x_i - \mu_{ij})^2}{2\sigma_{ij}^2}} \quad (7)$$

Di mana :

- $P$  : Peluang
- $X_i$  : Atribut ke  $i$
- $x_i$  : Nilai atribut ke  $i$
- $Y$  : Kelas yang dicari
- $y_i$  : Sub kelas  $Y$  yang dicari
- $\mu$  : *mean*, menyatakan rata – rata dari seluruh atribut
- $\sigma$  : *Deviasi standar*, menyatakan varian dari seluruh atribut.

Alur dari metode *Naive Bayes* dapat dilihat pada gambar 2 sebagai berikut:



Gambar 1. Alur Metode Naive Bayes

Adapun keterangan dari gambar 2 di atas sebagai berikut:

1. Baca *data training*
2. Hitung Jumlah dan probabilitas, namun apabila data numerik maka:
  - a. Cari nilai *mean* dan standar deviasi dari masing-masing parameter yang merupakan data numerik.

Adapun persamaan yang digunakan untuk menghitung nilai rata – rata hitung (*mean*) dapat dilihat sebagai berikut :

$$\mu = \frac{\sum_{i=1}^n x_i}{n} \quad (8)$$

atau

$$\mu = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} \quad (9)$$

di mana :

$\mu$  : rata – rata hitung (*mean*)

$x_i$  : nilai sample ke -*i*

$n$  : jumlah sampel

Dan persamaan untuk menghitung nilai simpangan baku (standar deviasi) dapat dilihat sebagai berikut:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n - 1}} \quad (10)$$

di mana :

$\sigma$  : standar deviasi

$x_i$  : nilai  $x$  ke  $-i$

$\mu$  : rata-rata hitung

$n$  : jumlah sampel

- b. Cari nilai probabilitik dengan cara menghitung jumlah data yang sesuai dari kategori yang sama dibagi dengan jumlah data pada kategori tersebut.
3. Mendapatkan nilai dalam tabel *mean*, standard deviasi dan probabilitas.
4. Solusi kemudian dihasilkan.

## 2. METODE PENELITIAN

Adapun metode penelitian yang diterapkan pada penelitian ini adalah sebagai berikut:

1. Analisis Masalah dan Studi Literatur  
Tahap ini adalah langkah awal untuk menentukan rumusan masalah dari penelitian. Dalam hal ini mengamati permasalahan yang berhubungan dengan faktor-faktor penyebab besarnya pemakaian listrik rumah tangga. Permasalahan-permasalahan yang ada, selanjutnya dianalisa untuk mengetahui bagaimana cara penyelesaian terhadap masalah tersebut dan menentukan ruang lingkup permasalahan yang akan diteliti. Mempelajari dasar teori dari berbagai literatur mengenai penerapan metode *Naïve Bayes*, konsep dan teori data mining dan penggunaan listrik rumah tangga, melalui jurnal-jurnal dan agar mendapatkan dasar pengetahuan untuk melakukan penelitian selanjutnya.
2. Mengumpulkan Data  
Prosedur sistematis yang digunakan untuk mengumpulkan data yaitu dengan mengajukan pertanyaan-pertanyaan kepada responden. Hasil kuesioner yang didapatkan akan digunakan untuk bahan penganalisisan data terhadap metode *Naïve Bayes*. Setelah data dikumpulkan dilakukan analisa data untuk menyesuaikan proses data yang akan diolah pada metode *Naïve Bayes*.
3. Implementasi dan Pengujian  
Sesuai dengan pengolahan data maka pada tahap implementasi adalah tentang bagaimana pengolahan datanya diterapkan dalam sebuah tools. Tools yang akan digunakan dalam implementasi penelitian ini adalah dengan menggunakan *Software Weka*. Selanjutnya Pengujian dilakukan untuk mengetahui apakah penelitian yang dilakukan telah sesuai dengan tujuan yang diharapkan.

## 3. HASIL DAN PEMBAHASAN

### 3.1. Penerapan Metode Naive Bayes

Naive Bayes didasarkan pada asumsi penyederhanaan bahwa nilai atribut secara kondisional saling bebas jika diberikan nilai output. Dengan kata lain, diberikan nilai output, probabilitas mengamati secara bersama adalah produk dari probabilitas individu. *Keuntungan penggunaan Naive Bayes adalah bahwa metode ini hanya membutuhkan jumlah data pelatihan (Training Data) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. Dalam metode Naive Bayes data String yang bersifat konstan dibedakan dengan data numerik yang bersifat kontinyu, perbedaan ini akan terlihat pada saat menentukan*

nilai probabilitas setiap kriteria baik itu kriteria dengan nilai data string maupun kriteria dengan nilai data numerik. Adapun penerapan metode Naive Bayes sebagai berikut.

### 3.1.1. Baca Data Training

Untuk menentukan data yang nantinya akan dianalisis dengan metode *Naive Bayes* maka langkah pertama yang dilakukan adalah membaca data latih. Adapun data latih yang digunakan dapat dilihat pada tabel 1 berikut[12]:

Tabel 1. Data Training

No	Jumlah Tanggungan Keluarga	Luas Rumah	Pendapatan/ bulan	Daya Listrik	Perlengkapan Yang Dimiliki	Penggunaan Listrik
1	Banyak	Besar	Besar	Sedang	Banyak	Tinggi
2	Banyak	Besar	Besar	Sedang	Banyak	Tinggi
3	Banyak	Besar	Besar	Sedang	Banyak	Tinggi
4	Banyak	Besar	Besar	Sedang	Banyak	Tinggi
5	Banyak	Besar	Besar	Sedang	Banyak	Tinggi
6	Kecil	Standar	Besar	Rendah	Sedang	Sedang
7	Kecil	Besar	Besar	Sedang	Sedang	Sedang
8	Kecil	Kecil	Besar	Sedang	Sedang	Sedang
9	Sedang	Besar	Besar	Sedang	Banyak	Sedang
10	Sedang	Besar	Besar	Sedang	Banyak	Sedang
11	Sedang	Standar	Besar	Sedang	Banyak	Sedang
12	Sedang	Standar	Besar	Sedang	Banyak	Sedang
13	Sedang	Standar	Besar	Sedang	Banyak	Tinggi
14	Sedang	Standar	Besar	Sedang	Banyak	Tinggi
15	Sedang	Standar	Besar	Sedang	Banyak	Tinggi
60	Banyak	Standar	Kecil	Sedang	Banyak	Sedang

### 3.1.2. Kriteria dan Probabilitas

Adapun nilai probabilitas setiap kriteria didapatkan dari data latih pada tabel 1. Adapun nilai probabilitas setiap kriteria sebagai berikut.

#### 3.1.2.1. Probabilitas Kriteria Jumlah Tanggungan

Berdasarkan data penggunaan listrik rumah tangga pada tabel 1 diketahui jumlah data latih (*data training*) adalah sebanyak 60 data, di mana dari 60 data tersebut terdapat 3 data rumah tangga dengan jumlah tanggungan banyak dan penggunaan listrik rendah, 2 data rumah tangga dengan jumlah tanggungan banyak dan penggunaan listrik sedang, 5 data rumah tangga dengan jumlah tanggungan banyak dan penggunaan listrik tinggi. 21 data rumah tangga dengan jumlah tanggungan sedang dan penggunaan listrik rendah, 4 data rumah tangga dengan jumlah tanggungan sedang dan penggunaan listrik sedang, 0 data rumah tangga dengan jumlah tanggungan sedang dan penggunaan listrik tinggi. 9 data rumah tangga dengan jumlah tanggungan sedikit dan penggunaan listrik rendah, 13 data rumah tangga dengan jumlah tanggungan sedikit dan penggunaan listrik sedang, 3 data rumah tangga dengan jumlah tanggungan sedikit dan penggunaan listrik tinggi. Probabilitas kriteria jumlah tanggungan dapat dilihat pada tabel 2.

Tabel 2. Probabilitas Kriteria Jumlah Tanggungan

Jumlah Tanggungan	Jumlah Kejadian “Penggunaan Listrik”			Probabilitas		
	Rendah	Sedang	Tinggi	Rendah	Sedang	Tinggi
Banyak	3	2	5	0.09	0.11	0.63
Sedang	21	4	0	0.64	0.21	0.00
Sedikit	9	13	3	0.27	0.68	0.38
<b>Jumlah</b>	33	19	8	0.55	0.32	0.13

### 3.1.2.2. Probabilitas Kriteria Luas Rumah

Pada kriteria luas rumah dapat diketahui dari 60 data terdapat 7 data rumah tangga dengan luas rumah besar dan penggunaan listrik rendah, 5 data rumah tangga dengan luas rumah besar dan penggunaan listrik sedang, 5 data rumah tangga dengan luas rumah besar dan penggunaan listrik tinggi. 5 data rumah tangga dengan luas rumah standar dan penggunaan listrik rendah, 11 data rumah tangga dengan luas rumah standar dan penggunaan listrik sedang, 3 data rumah tangga dengan luas rumah standar dan penggunaan listrik tinggi. 21 data rumah tangga dengan luas rumah kecil dan penggunaan listrik rendah, 3 data rumah tangga dengan luas rumah kecil dan penggunaan listrik sedang, 0 data rumah tangga dengan luas rumah kecil dan penggunaan listrik tinggi. Probabilitas kriteria luas rumah dapat dilihat pada tabel 3.

Tabel 3. Probabilitas Kriteria Luas Rumah

Luas Rumah	Jumlah Kejadian “Penggunaan Listrik”			Probabilitas		
	Rendah	Sedang	Tinggi	Rendah	Sedang	Tinggi
Besar	7	5	5	0.21	0.26	0.63
Standar	5	11	3	0.15	0.58	0.38
Kecil	21	3	0	0.64	0.16	0.00
<b>Jumlah</b>	33	19	8	0.55	0.32	0.13

### 3.1.2.3. Probabilitas Kriteria Pendapatan

Pada kriteria pendapatan dapat diketahui dari 60 data terdapat 0 data rumah tangga dengan pendapatan besar dan penggunaan listrik rendah, 7 data rumah tangga dengan pendapatan besar dan penggunaan listrik sedang, 8 data rumah tangga dengan pendapatan besar dan penggunaan listrik tinggi. 0 data rumah tangga dengan pendapatan sedang dan penggunaan listrik rendah, 0 data rumah tangga dengan pendapatan sedang dan penggunaan listrik sedang, 0 data rumah tangga dengan pendapatan sedang dan penggunaan listrik tinggi. 33 data rumah tangga dengan pendapatan kecil dan penggunaan listrik rendah, 12 data rumah tangga dengan pendapatan kecil dan penggunaan listrik sedang, 0 data rumah tangga dengan pendapatan kecil dan penggunaan listrik tinggi. Probabilitas kriteria pendapatan dapat dilihat pada tabel 4.

Tabel 4. Probabilitas Pendapatan

Pendapatan	Jumlah Kejadian “Penggunaan Listrik”			Probabilitas		
	Rendah	Sedang	Tinggi	Rendah	Sedang	Tinggi
Besar	0	7	8	0.00	0.37	1.00
Sedang	0	0	0	0.00	0.00	0.00
Kecil	33	12	0	1.00	0.63	0.00
<b>Jumlah</b>	33	19	8	0.55	0.32	0.13

3.1.2.4. Probabilitas Daya Listrik

Pada kriteria daya listrik dapat diketahui dari 60 data terdapat 0 data rumah tangga dengan daya listrik tinggi dan penggunaan listrik rendah, 7 data rumah tangga dengan daya listrik tinggi dan penggunaan listrik sedang, 8 data rumah tangga dengan daya listrik tinggi dan penggunaan listrik tinggi. 0 data rumah tangga dengan daya listrik sedang dan penggunaan listrik rendah, 0 data rumah tangga dengan daya listrik sedang dan penggunaan listrik sedang, 0 data rumah tangga dengan daya listrik sedang dan penggunaan listrik tinggi. 33 data rumah tangga dengan daya listrik rendah dan penggunaan listrik rendah, 12 data rumah tangga dengan daya listrik rendah dan penggunaan listrik sedang, 0 data rumah tangga dengan daya listrik rendah dan penggunaan listrik tinggi. Probabilitas kriteria daya listrik dapat dilihat pada tabel 5.

Tabel 5. Probabilitas Daya Listrik

Daya Listrik	Jumlah Kejadian “Penggunaan Listrik”			Probabilitas		
	Rendah	Sedang	Tinggi	Rendah	Sedang	Tinggi
Tinggi	0	0	0	0.00	0.00	0.00
Sedang	21	17	8	0.64	0.89	1.00
Rendah	12	2	0	0.36	0.11	0.00
Jumlah	33	19	8	0.55	0.32	0.13

3.1.2.5. Probabilitas Perlengkapan

Pada kriteria perlengkapan dapat diketahui dari 60 data terdapat 0 data rumah tangga dengan perlengkapan tinggi dan penggunaan listrik rendah, 7 data rumah tangga dengan perlengkapan banyak dan penggunaan listrik sedang, 8 data rumah tangga dengan perlengkapan banyak dan penggunaan listrik banyak. 0 data rumah tangga dengan perlengkapan sedang dan penggunaan listrik rendah, 0 data rumah tangga dengan perlengkapan sedang dan penggunaan listrik sedang, 0 data rumah tangga dengan perlengkapan sedang dan penggunaan listrik tinggi. 33 data rumah tangga dengan perlengkapan sedikit dan penggunaan listrik rendah, 12 data rumah tangga dengan perlengkapan sedikit dan penggunaan listrik sedang, 0 data rumah tangga dengan perlengkapan sedikit dan penggunaan listrik tinggi. Probabilitas kriteria perlengkapan dapat dilihat pada tabel 6.

Tabel 6. Probabilitas Perlengkapan

Perlengkapan	Jumlah Kejadian “Penggunaan Listrik”			Probabilitas		
	Rendah	Sedang	Tinggi	Rendah	Sedang	Tinggi
Banyak	4	12	8	0.12	0.63	1.00
Sedang	3	3	0	0.09	0.16	0.00
Sedikit	26	4	0	0.79	0.21	0.00
Jumlah	33	19	8	0.55	0.32	0.13

3.1.2.6. Probabilitas Penggunaan Listrik

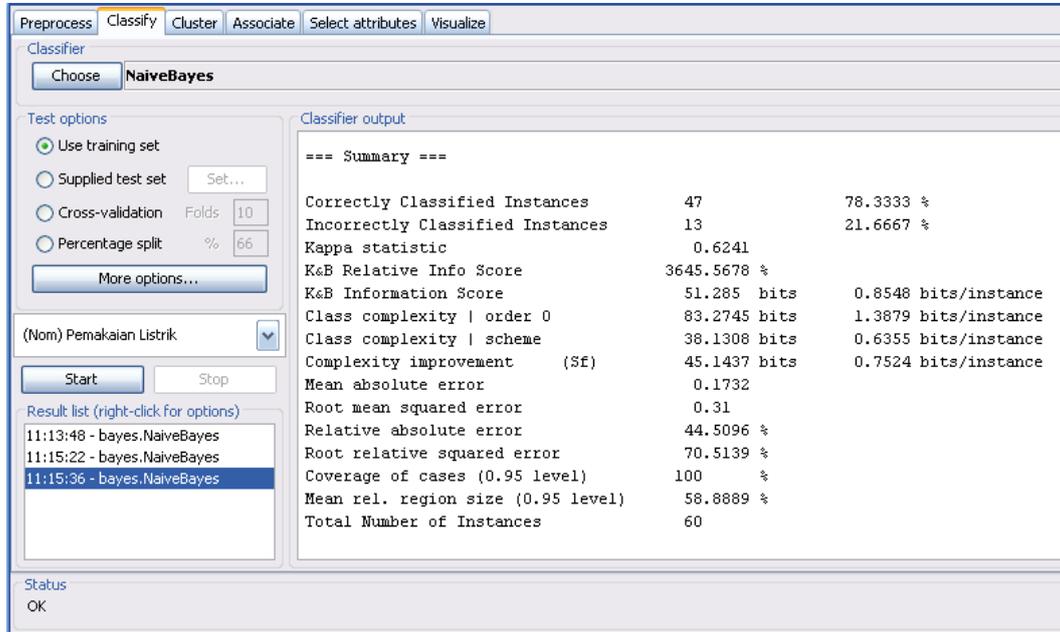
Berdasarkan tabel 1 diketahui dari 60 data penggunaan listrik terdapat 33 data rumah tangga dengan penggunaan listrik rendah, 19 data rumah tangga dengan penggunaan listrik sedang, 8 data rumah tangga dengan penggunaan listrik tinggi. Probabilitas kriteria penggunaan listrik dapat dilihat pada tabel 7.

Tabel 7. Probabilitas Penggunaan Listrik

Jumlah Kejadian “Penggunaan Listrik”			Probabilitas		
Rendah	Sedang	Tinggi	Rendah	Sedang	Tinggi
33	19	8	0.55	0.32	0.13

### 3.2. Pengujian Metode Naive Bayes

Dari nilai probabilitas di atas akan diuji data sebanyak 60 data dan diselesaikan dengan menggunakan tools weka sehingga dihasilkan hasil klasifikasi penggunaan listrik seperti pada gambar 3 berikut:



Gambar 3. Hasil Klasifikasi Metode Naive Bayes

Berdasarkan gambar 3 di atas dapat dilihat persentase untuk *Correctly Classified Instance* adalah sebesar 78,3333% sementara persentase untuk *Incorrectly Classified Instance* adalah sebesar 21,6667%. Di mana dari 60 data penggunaan listrik rumah tangga, ada sebanyak 47 data penggunaan listrik rumah tangga berhasil diklasifikasikan dengan benar dan sebanyak 13 data penggunaan listrik rumah tangga tidak berhasil diklasifikasikan dengan benar.

## 4. KESIMPULAN

Berdasarkan penelitian tentang memprediksi penggunaan listrik rumah tangga dapat ditarik beberapa kesimpulan sebagai berikut :

1. Metode *Naive Bayes* memanfaatkan data *training* untuk menghasilkan probabilitas setiap kriteria untuk *class* yang berbeda, sehingga nilai-nilai probabilitas dari kriteria tersebut dapat dioptimalkan untuk memprediksi penggunaan listrik berdasarkan proses klasifikasi yang dilakukan oleh metode *Naive Bayes* itu sendiri.
2. Berdasarkan data rumah tangga yang dijadikan data *training*, metode *Naive Bayes* berhasil mengklasifikasikan 47 data dari 60 data yang diuji. Sehingga metode *Naive Bayes* berhasil memprediksi besarnya penggunaan listrik rumah tangga dengan persentase keakuratan sebesar 78,3333%.

## 5. SARAN

Adapun saran yang dapat diberikan mengenai penelitian tentang memprediksi penggunaan listrik rumah tangga yaitu pengujian sebaiknya dilakukan dengan menggunakan metode lain untuk melihat metode mana yang lebih akurat dalam memprediksi penggunaan listrik rumah tangga

## DAFTAR PUSTAKA

- [1] Harifuddin, 2007, Estimasi Kebutuhan Daya Listrik Sulawesi Selatan Sampai Tahun 2017, *Media Elektrik*, No. 2 Vol. 2, Hal 14-22.
  - [2] Chahaya, S., I., 2005, Pengetahuan, Sikap dan Tindakan Masyarakat Dalam Upaya Menghemat Pemakaian Energi Listrik Di Perumahan Nasional (Perumnas) Helvetia Kecamatan Medan Helvetia Kota Medan, *Jurnal Komunikasi Penelitian*, Vol. 17, No.4, Hal 60-65.
  - [3] Manjusha K. K., Sankaranarayanan, K., Seena P., 2014, Prediction of Different Dermatological Conditions Using Naive Bayesian Classification, *International Journal of Advanced Research in Computer Science and Software Engineering*, Vol 4, No 1, Hal 864-868.
  - [4] Ting, S. L., Ip, W. H., Tsang, A. H.C., 2011, Is Naive Bayes a Good Classifier for Document Classification?, *International Journal of Software Engineering and Its Applications*, Vol. 5, No. 3, Hal 37-46.
  - [5] Taruna R., S., Hiranwal, S., 2013, Enhanced Naive Bayes Algorithm for Intrusion Detection in Data Mining, *International Journal of Computer Science and Information Technologies*, Vol.6, No. 4, Hal 960-962.
  - [6] Mabur, A. G., Lubis, R., 2012, Penerapan Data Mining untuk Memprediksi Kriteria Nasabah Kredit, *Jurnal Komputer dan Informatika (KOMPUTA)*, Vol.1, No. 1, Hal 53-57.
  - [7] Mujiasih, S., 2011, Pemanfaatan Data Mining Untuk Prakiraan Cuaca, *Jurnal Meteorologi dan Geofisika*, Vol. 12, No. 2, Hal 189 - 195.
  - [8] Ridwan, M., Suyono, H., Sarosa, M., 2013, Penerapan Data Mining untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier, *Jurnal EECCIS*, Vol 1, No. 7, Hal. 59-64.
  - [9] Patil, T. R., Sherekar, M. S., 2013, Performance Analysis of Naive Bayes and J48 Classification Algorithm for Data Classification, *International Journal of Computer Science and Applications*, Vol. 6, No. 2, Hal 256-261.
  - [10] Bustami., 2013, Penerapan Algoritma Naive Bayes Untuk Mengklasifikasi Data Nasabah Asuransi, *TECHSI : Jurnal Penelitian Teknik Informatika*, Vol. 3, No.2, Hal. 127-146.
  - [11] Pattekari, S. A., Parveen, A., 2012, Prediction System for Heart Disease Using Naive Bayes, *International Journal of Advanced Computer and Mathematical Sciences*, ISSN 2230-9624, Vol. 3, No 3, Hal 290-294.
  - [12] Nasari, F., 2014, Analisa Faktor Penyebab Tingginya Pemakaian Listrik Rumah Tangga Menggunakan Algoritma C 4.5 (Studi Kasus di Kelurahan Tanjung Mulia), *Tesis*, Program Pasca Sarjana Ilmu Komputer, Universitas Putra Indonesia (YPTK), Padang.
-