



## Klasifikasi Tingkat Penjualan Air Minum Isi Ulang Harian Menggunakan *Naive Bayes*

Angga Lisdiyanto <sup>a</sup>, Winarti <sup>b</sup>, Catur Lega Wibisono <sup>c</sup>

<sup>a</sup> Teknik Informatika, Fakultas Ilmu Komputer, Universitas Nahdlatul Ulama Sidoarjo

<sup>b</sup> Teknik Informatika Fakultas Teknik, Universitas Darul Ulum, Jombang, Indonesia

<sup>c</sup> Teknik Informatika, Fakultas Ilmu Komputer, Universitas Nahdlatul Ulama Sidoarjo

email: <sup>a</sup> [angga.ti@unusida.ac.id](mailto:angga.ti@unusida.ac.id), <sup>b</sup> [winarti.ti@undar.ac.id](mailto:winarti.ti@undar.ac.id), <sup>c</sup> [catur.si@unusida.ac.id](mailto:catur.si@unusida.ac.id)

### INFORMASI ARTIKEL

**Sejarah artikel:**  
Accepted 20 April 2022

**Kata kunci:**  
Prediksi  
Penjualan  
Naive Bayes  
Air Minum

### A B S T R A K

Perkembangan teknologi informasi pada zaman ini berkembang dengan sangat cepat dan pesat. Salah satu alat teknologi yang dimaksud adalah komputer. Dunia usaha yang terus berubah dengan cepat, serta banyaknya persaingan pengusaha yang semakin banyak menjadikan pelaku usaha harus memiliki strategi penjualan. Kegiatan prediksi masa depan merupakan salah satu usaha perusahaan sebagai dasar pengambilan keputusan strategis dalam kelangsungan usaha. Tujuan penelitian ini adalah membangun sebuah sistem yang dapat membantu depo untuk merencanakan strategi penjualan masa yang akan datang. Dimana nantinya akan memunculkan sebuah prediksi penjualan lebih atau kurang berdasarkan perhitungan yang ditentukan. Dalam sistem ini metode Naive Bayes digunakan sebagai dasar untuk melakukan perhitungan prediksi penjualan yang diinputkan, sehingga memunculkan suatu hasil prediksi. uji coba prediksi penjualan air minum isi ulang menggunakan Naive Bayes. Sistem prediksi ini dapat membantu untuk menyiapkan air yang cukup untuk melayani penjualan "Lebih".

### 1. Pendahuluan

Banyak algoritma klasifikasi yang terkenal, pada penelitian yang sudah ada sebelumnya para peneliti menggunakan berbagai macam metode, diantaranya Naive bayes, Apriori, K-NN dan lain-lain, dengan kelebihan dan kekurangan masing-masing metode tersebut. Namun pada penelitian ini akan menggunakan metode klasifikasi Decision Tree C4.5 guna untuk mencoba mengembangkan penelitian sebelumnya dengan objek berbeda, yaitu prediksi kelulusan mahasiswa [3]. Dalam klasifikasi, variabel target-nya merupakan kategorikal. Model data mining memeriksa set record yang besar, tiap record mempunyai informasi variabel target dan set input atau variabel predictor. Kelebihan dari metode decision tree adalah daerah pengambilan keputusan yang sebelumnya kompleks dan sangat global, dapat diubah menjadi lebih simple dan spesifik. Decision Tree adalah sebuah struktur pohon, dimana setiap node pohon merepresentasikan atribut yang telah diuji, setiap cabang merupakan suatu pembagian hasil uji, dan node daun (leaf) merepresentasikan kelompok kelas tertentu. Level node teratas dari sebuah Decision Tree adalah node akar (root) yang biasanya berupa atribut yang paling memiliki pengaruh terbesar pada suatu kelas tertentu. Decision Tree melakukan strategi pencarian secara top-down untuk solusinya.

### 2. State of the Art

Penggunaan teknik data mining diharapkan dapat membantu percepatan proses pengambilan keputusan, memungkinkan perusahaan untuk mengelola informasi yang terkandung didalam data nasabah dan menjadi sebuah pengetahuan (knowledge) yang baru untuk menentukan keputusan yang dibutuhkan. Oleh karena itu, untuk benar-benar mengetahui dan menetapkan jenis vitamin ini LAKU atau TIDAK LAKU perlu dilakukan forecasting (prediksi/estimasi) sehingga jenis vitamin tersebut akan sesuai dengan jenis vitamin yang diminati oleh pelanggan dari apotek tersebut. Setelah jenis obat ditetapkan, kemudian diestimasi berapa banyak permintaan jenis vitamin tersebut pada bulan berikutnya dengan menggunakan algoritma Naive Bayes. Data Mining adalah teknik untuk memproses dan mengekstraksi data besar menjadi informasi yang bisa membentuk data baru [1].

Penelitian sebelumnya juga dapat digunakan untuk memprediksi nasabah sebuah perusahaan asuransi untuk mengetahui lancar, kurang lancar atau tidak lancarnya nasabah tersebut dalam pembayaran pinjaman tersebut [2].

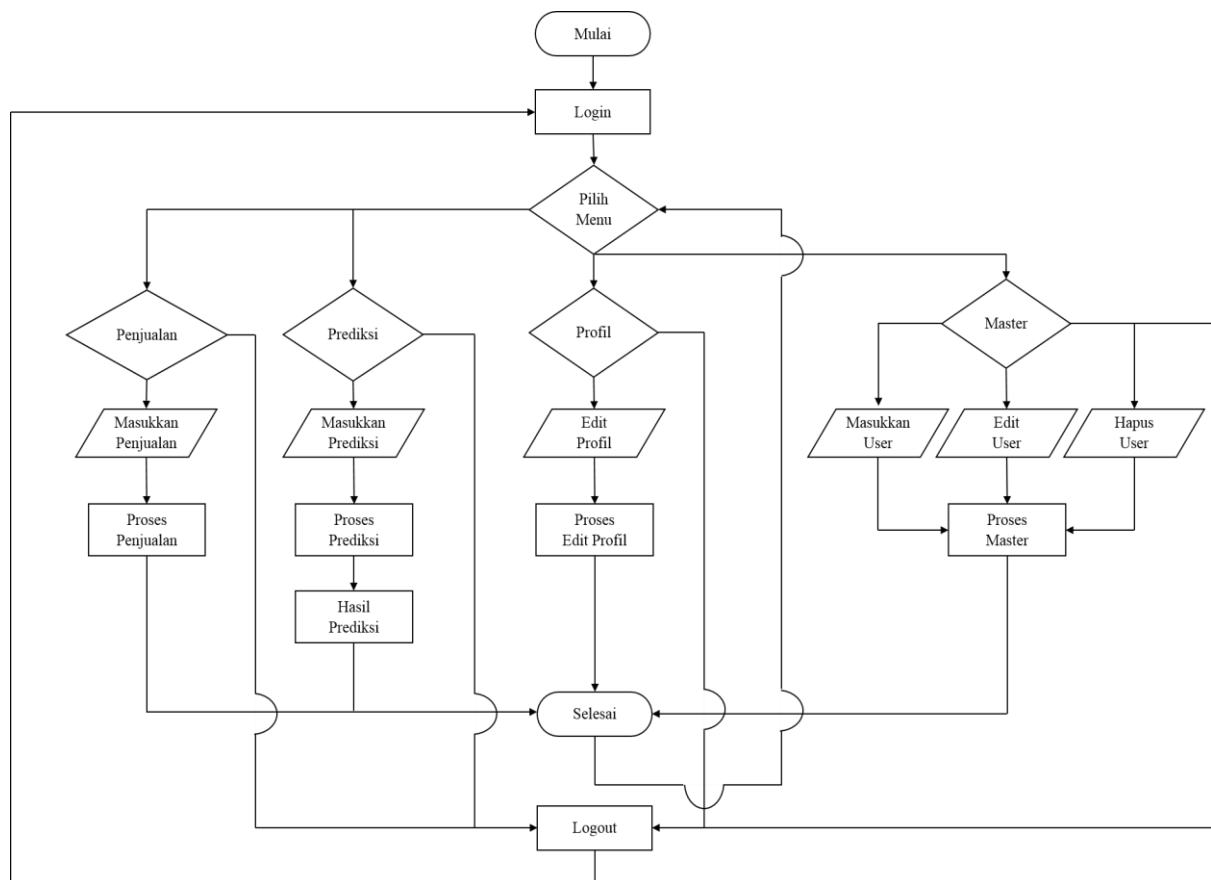
Berawal dari beberapa disiplin ilmu yang ada, data mining memiliki tujuan untuk memperbaiki teknik tradisional yang dapat menghasilkan sesuatu yang jauh lebih informatif sehingga bisa menangani: Jumlah data yang sangat besar, Dimensi data yang tinggi dan Data yang heterogen dan berbeda sifat satu dengan yang lainnya [3].

### 3. Method

Data yang diperoleh dengan melakukan penelitian pada tanggal 01 Juli 2020 sampai tanggal 04 Juli 2020 adalah berjumlah 122 hari dari tanggal 01 Maret 2020 hingga 01 Juli 2020.

Tahap ini melakukan perancangan desain sistem prediksi penjualan dengan menggunakan tabel setiap data yang diperlukan untuk membantu pembuatan sistem prediksi penjualan.

Berikut adalah perancangan sistem prediksi :



Gambar 1 Flowchart Sistem Prediksi

Dalam data training dan data uji coba ada beberapa atribut yang dipakai yaitu tanggal, hari, jadwal, cuaca, jumlah isi, dan jumlah kirim. Sedangkan klas adalah penjualan.

1. Hari  
Tanggal adalah dasar dari semua atribut, jika tidak ada tanggal maka tidak akan bisa menentukan jadwal, jumlah dan penjualan.
2. Jadwal  
Jadwal berfungsi untuk membedakan dari setiap tanggal serta membedakan untuk perhitungan jumlah dalam menentukan hasil penjualan
3. Cuaca  
Cuaca terdiri dari hujan dan tidak hujan
4. Jumlah Isi  
Adalah jumlah pengisian air galon dalam 1 hari
5. Jumlah Kirim  
Adalah jumlah pengiriman air galon ke konsumen dalam 1 hari
6. Penjualan  
Atribut penjualan adalah klas yang terdiri dari kurang dan lebih

Berdasarkan atribut jumlah isi dan jumlah kirim, maka kriteria tersebut menghasilkan rumus sebagai berikut :

- a.  $\text{Jumlah\_Isi} = (\text{Jumlah Isi} - \text{Jumlah Kirim}) * 4000$
- b.  $\text{Jumlah\_Kirim} = \text{Jumlah Kirim} * 5000$
- c.  $\text{Jumlah} = \text{Jumlah\_Isi} + \text{Jumlah Kirim}$

Tabel 1 Data Training

Tanggal	Hari	Jadwal	Cuaca	Jumlah Isi	Jumlah Kirim	Penjualan
01/03/2020	Minggu	Tanggal Merah	Tidak Hujan	97	54	Lebih
02/03/2020	Senin	Bukan Tanggal Merah	Jika Hujan	78	46	Lebih
03/03/2020	Selasa	Bukan Tanggal Merah	Tidak Hujan	64	40	Kurang
04/03/2020	Rabu	Bukan Tanggal Merah	Tidak Hujan	77	51	Lebih
05/03/2020	Kamis	Bukan Tanggal Merah	Tidak Hujan	80	57	Lebih
06/03/2020	Jumat	Bukan Tanggal Merah	Tidak Hujan	64	49	Kurang
07/03/2020	Sabtu	Bukan Tanggal Merah	Tidak Hujan	76	63	Lebih
08/03/2020	Minggu	Tanggal Merah	Jika Hujan	70	37	Kurang
09/03/2020	Senin	Bukan Tanggal Merah	Tidak Hujan	76	42	Kurang
10/03/2020	Selasa	Bukan Tanggal Merah	Tidak Hujan	116	68	Lebih
11/03/2020	Rabu	Bukan Tanggal Merah	Tidak Hujan	83	52	Lebih
12/03/2020	Kamis	Bukan Tanggal Merah	Tidak Hujan	66	44	Kurang

Pada tabel 1 adalah data yang digunakan sebagai data training untuk melakukan proses prediksi penjualan, adapun nilai konversi setiap kolom dalam tabel 1 sebagai berikut :

- a. Konversi data hari
  - 1) Senin = 0.1
  - 2) Selasa = 0.2
  - 3) Rabu = 0.3
  - 4) Kamis = 0.4
  - 5) Jumat = 0.5
  - 6) Sabtu = 0.6
  - 7) Minggu = 0.7
- b. Konversi data jadwal
  - 1) Tanggal merah = 0.1
  - 2) Bukan tanggal merah = 0.2
- c. Konversi data cuaca
  - 1) Jika hujan = 0.1
  - 2) Tidak hujan = 0.2

## d. Konversi jumlah isi

- 1) 0 – 20 = 0.1
- 2) 21 – 40 = 0.2
- 3) 41 – 60 = 0.3
- 4) 61 – 80 = 0.4
- 5) 81 – 100 = 0.5
- 6) 101 – 120 = 0.6
- 7) 121 – 140 = 0.7
- 8) 141 – 160 = 0.8
- 9) > 160 = 0.9

## e. Konversi jumlah kirim

- 1) 0 – 20 = 0.1
- 2) 21 – 40 = 0.2
- 3) 41 – 60 = 0.3
- 4) 61 – 80 = 0.4
- 5) 81 – 100 = 0.5
- 6) 101 – 120 = 0.6
- 7) 121 – 140 = 0.7
- 8) 141 – 160 = 0.8
- 9) > 160 = 0.9

## f. Konversi data penjualan

- 1) Lebih = 0.1
- 2) Kurang = 0.2

Konversi digunakan untuk proses penambahan data penjualan dan melakukan proses perhitungan prediksi dalam sistem aplikasi.

Berikut kriteria untuk menghasilkan penjualan lebih atau kurang :

1. Jika tanggal tersebut menyatakan tanggal merah dan ketika perhitungan jumlah  $\geq 340000$ , maka pada tanggal tersebut akan dinyatakan penghasilan lebih.
2. Jika tanggal tersebut menyatakan tanggal merah dan ketika perhitungan jumlah  $< 340000$ , maka pada tanggal tersebut akan dinyatakan penghasilan kurang.
3. Jika tanggal tersebut menyatakan bukan tanggal merah dan ketika perhitungan jumlah  $\geq 350000$ , maka pada tanggal tersebut akan dinyatakan penghasilan lebih.
4. Jika tanggal tersebut menyatakan bukan tanggal merah dan ketika perhitungan jumlah  $< 350000$ , maka pada tanggal tersebut akan dinyatakan penghasilan kurang.

#### 4. Hasil and Pembahasan

Tabel 2 Data Prediksi

Tanggal	Hari	Jadwal	Cuaca
16/03/2020	Senin	Bukan Tanggal Merah	Tidak Hujan

Berdasarkan tabel 2 tanggal "16/03/2020", hari "Senin", jadwal "Bukan Tanggal Merah" dan cuaca "Tidak Hujan" ketika menggunakan metode Naive Bayes, maka akan mnghasilkan seperti berikut :

$P(\text{Hari Senin} \mid \text{Penjualan} = \text{Lebih})$	: 1 / 7	= 0.14
$P(\text{Hari Senin} \mid \text{Penjualan} = \text{Kurang})$	: 1 / 5	= 0.2
$P(\text{Jadwal Bukan Tanggal Merah} \mid \text{Penjualan} = \text{Lebih})$	: 6 / 7	= 0.85
$P(\text{Jadwal Bukan Tanggal Merah} \mid \text{Penjualan} = \text{Kurang})$	: 4 / 5	= 0.8
$P(\text{Cuaca Tidak Hujan} \mid \text{Penjualan} = \text{Lebih})$	: 6 / 7	= 0.85
$P(\text{Cuaca Tidak Hujan} \mid \text{Penjualan} = \text{Kurang})$	: 4 / 5	= 0.8

---

$P(\text{Hari Senin} \mid \text{Penjualan} = \text{Lebih})^*$		
$P(\text{Jadwal Bukan Tanggal Merah} \mid \text{Penjualan} = \text{Lebih})^*$	:	$0.14 * 0.85 * 0.85 = 0.1$
$P(\text{Cuaca Tidak Hujan} \mid \text{Penjualan} = \text{Lebih})$		
$P(\text{Hari Senin} \mid \text{Penjualan} = \text{Kurang})^*$		
$P(\text{Jadwal Bukan Tanggal Merah} \mid \text{Penjualan} = \text{Kurang})^*$	:	$0.2 * 0.8 * 0.8 = 0.13$
$P(\text{Cuaca Tidak Hujan} \mid \text{Penjualan} = \text{Kurang})$		
$P(X \mid \text{Penjualan} = \text{Lebih})$	:	$0.1 * (7/12) = 0.06$
$P(X \mid \text{Penjualan} = \text{Kurang})$	:	$0.13 * (5/12) = 0.05$

Dalam perhitungan di atas, dapat dilihat bahwa  $P(X \mid \text{Penjualan} = \text{Lebih})$  bernilai lebih besar daripada nilai  $P(X \mid \text{Penjualan} = \text{Kurang})$ . Perhitungan prediksi menghasilkan penjualan “Lebih” dengan nilai 0.6, maka data prediksi kolom penjualan maka ditampilkan hasil klasifikasi “Lebih”, seperti pada tabel 3.

Tabel 3 Hasil Data Prediksi

Tanggal	Hari	Jadwal	Cuaca	Penjualan
16/03/2020	Senin	Bukan Tanggal Merah	Tidak Hujan	Lebih

Berdasarkan tabel 3 tanggal “16/03/2020”, hari “Senin”, jadwal “Bukan Tanggal Merah” dan cuaca “Tidak Hujan” ketika menggunakan metode Naive Bayes maka penjualan adalah “Lebih”.

Berikut adalah uji coba prediksi dengan beberapa kombinasi jenis jadwal dan jenis cuaca :

1. Prediksi ke-1 dilakukan dengan memilih pada tanggal “01 Agustus 2020” dengan jadwal “Tanggal Merah” serta cuaca “Jika Hujan”. Hasil prediksi penjualan pada tanggal “01 Agustus 2020” menyatakan hasil penjualan adalah “Lebih” dengan jumlah pengisian 94 dan jumlah pengiriman 58.
  2. Prediksi ke-2 dilakukan dengan memilih pada tanggal “01 Agustus 2020” dengan jadwal “Tanggal Merah” serta cuaca “Tidak Hujan”. Hasil prediksi penjualan pada tanggal “01 Agustus 2020” menghasilkan pernyataan bahwa hasil penjualan adalah “Lebih” dengan banyaknya jumlah pengisian 94 dan banyaknya jumlah pengiriman 59.
  3. Prediksi ke-3 dilakukan dengan memilih pada tanggal “01 Agustus 2020” dengan jadwal “Bukan Tanggal Merah” serta cuaca “Jika Hujan”. Hasil prediksi penjualan pada tanggal “01 Agustus 2020” menyatakan hasil penjualan adalah “Lebih” dengan jumlah pengisian 92 dan jumlah pengiriman 57.
  4. Prediksi ke-4 dilakukan pada tanggal “01 Agustus 2020”, jadwal “Bukan Tanggal Merah” dan cuaca “Tidak Hujan”. Hasil prediksi penjualan pada tanggal “01 Agustus 2020” menyatakan hasil penjualan adalah “Lebih” dengan jumlah pengisian 92 dan jumlah pengiriman 58.
  5. Prediksi ke-5 dilakukan dengan memilih pada tanggal “05 Agustus 2020” dengan jadwal “Tanggal Merah” serta cuaca “Jika Hujan”. Hasil prediksi penjualan pada tanggal “05 Agustus 2020” menyatakan hasil penjualan adalah “Lebih” dengan jumlah pengisian 89 dan jumlah pengiriman 56.
  6. Prediksi ke-6 dilakukan dengan memilih pada tanggal “05 Agustus 2020” dengan jadwal “Tanggal Merah” serta cuaca “Tidak Hujan”. Hasil prediksi penjualan pada tanggal “05 Agustus 2020” menyatakan hasil penjualan adalah “Lebih” dengan jumlah pengisian 90 dan jumlah pengiriman 56.
  7. Prediksi ke-7 dilakukan dengan memilih pada tanggal “05 Agustus 2020” dengan jadwal “Bukan Tanggal Merah” serta cuaca “Jika Hujan”. Hasil prediksi penjualan pada tanggal “05 Agustus 2020” menyatakan hasil penjualan adalah “Kurang” dengan jumlah pengisian 66 dan jumlah pengiriman 44.
  8. Prediksi ke-8 dilakukan dengan memilih pada tanggal “05 Agustus 2020” dengan jadwal “Bukan Tanggal Merah” serta cuaca “Tidak Hujan”. Hasil prediksi penjualan pada tanggal “05 Agustus 2020” menyatakan hasil penjualan adalah “Lebih” dengan jumlah pengisian 88 dan jumlah pengiriman 55.
-

Sehingga didapatkan hasil prediksi seperti pada tabel berikut :

Tabel 4 Hasil Uji Coba Prediksi

Tanggal	Hari	Jadwal	Cuaca	Jumlah Isi	Jumlah Kirim	Klas Penjualan
01/05/2020	Sabtu	Tanggal Merah	Hujan	94	58	Lebih
01/05/2020	Sabtu	Tanggal Merah	Tidak Hujan	94	59	Lebih
01/05/2020	Sabtu	Bukan Tanggal Merah	Hujan	92	57	Lebih
01/05/2020	Sabtu	Bukan Tanggal Merah	Tidak Hujan	92	58	Lebih
05/05/2020	Rabu	Tanggal Merah	Hujan	89	55	Lebih
05/05/2020	Rabu	Tanggal Merah	Tidak Hujan	90	56	Lebih
05/05/2020	Rabu	Bukan Tanggal Merah	Hujan	66	44	Kurang

## 5. Kesimpulan

Berdasarkan hasil dan pembahasan uji coba prediksi penjualan air minum isi ulang menggunakan Naive Bayes, didapatkan kesimpulan bahwa sistem prediksi ini dapat membantu untuk menyiapkan air yang cukup untuk melayani penjualan "Lebih".

## 6. Referensi

- [1] Arif, B., Saruni, D. (2018). The Clasification Product Using Naïve Bayes Algorithm
- [2] Bustami. (2014). Penerapan Algoritma Naive Bayes Untuk Mengklasifikasi Data Nasabah Asuransi. Jurnal Informatika Vol. 8, No. 1, Januari.
- [3] Ratniasih, Ni Luh, 2019. Optimasi Data Mining Menggunakan Algoritma Naïve Bayes dan C4.5 Untuk Klasifikasi Kelulusan Mahasiswa. Jurnal Teknologi Informasi dan Komputer, Volume 5, Nomor 1, Januari 2019.