

## Penerapan Data Mining Pada Penjualan Makanan Dan Minuman Menggunakan Metode Algoritma *Naïve Bayes*

Fauzan Alghifari<sup>1</sup>, Didi Juardi<sup>2</sup>

<sup>1</sup>Universitas Singaperbangsa Karawang, Jl. H. S. Ronggowaluyo, Puseurjaya, Telukjambe Timur, Karawang, Jawa Barat 41361, Indonesia

<sup>2</sup>Universitas Singaperbangsa Karawang, Jl. H. S. Ronggowaluyo, Puseurjaya, Telukjambe Timur, Karawang, Jawa Barat 41361, Indonesia

### INFORMASI ARTIKEL

Sejarah Artikel:

Diterima Redaksi: 28 Juni 2021

Revisi Akhir: 01 Juli 2021

Diterbitkan Online: 10 September 2021

### KATA KUNCI

Data Mining

*Naïve Bayes*

Algoritma

### KORESPONDENSI

E-mail: [fauzan.alghifari17095@student.unsika.ac.id](mailto:fauzan.alghifari17095@student.unsika.ac.id)

### A B S T R A C T

Changes in human behavior against the times, this causes business actors, especially culinary businesses to innovate so that the business they run continues to survive. With this competence, it causes more problems, one of which is the emergence of startups that eventually become competitors in the business world in all fields. All business owners have an innovation by analyzing food and beverage sales to generate new opportunities which will then be used to identify future opportunities. This affects the income and expenses obtained by companies and restaurants. This research paper generates new knowledge, namely by using the Na Baye Bayes algorithm to generate an opportunity in the sale of food and beverages based on sales in the previous month. The results of the analysis process will be used for the benefit of restaurants in marketing optimization efforts. This new knowledge will also provide information such as the results of sales opportunities in the following months.

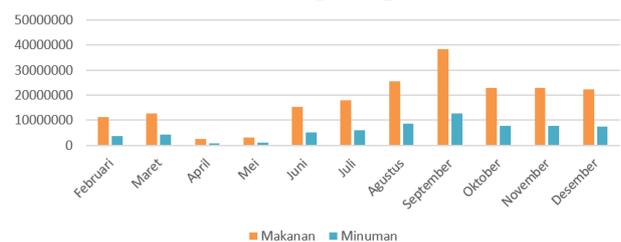
## 1. PENDAHULUAN

Pemanfaatan alat komunikasi belakangan ini sangat meningkat secara pesat, Dalam data mining, penambangan informasi digunakan untuk meningkatkan keuntungan dalam aktivitas jual beli. Salah satu manfaat algoritma *Naïve Bayes* dalam aktivitas jual beli dengan memprediksi peluang suatu class. Informasi ini digunakan untuk aktivitas jual beli supaya lebih efisien. Dengan informasi ini restoran Makan Barbeque Sepuasnya dapat mengetahui peluang pembeli terhadap makanan maupun minuman yang terdapat pada restoran tersebut.

Data yang sudah ada sebelumnya, masih banyak persoalan-persoalan yang terjadi di restoran Makan Barbeque Sepuasnya. Restoran ini hanya menjual satu menu saja. Pelanggan restoran ini banyak dari daerah Rawangun. Salah satu persoalan yang ada adalah skema penjualan yang kurang baik, melihat dari jumlah penjualan pada setiap bulannya. Hal ini tersebut banyaknya restoran lain yang membidangi yang sama. Oleh karena itu perlu adanya solusi untuk menghadapi persaingan dengan restoran lain yang membidangi yang sama. Berdasarkan data yang diperoleh dari restoran Makan Barbeque Sepuasnya, maka diperoleh data sebagai berikut :

Fauzan Alghifari

Data Penjualan  
Makan Barbeque Sepuasnya



Gambar 1 Data Penjualan Pada Restoran Makan Barbeque Sepuasnya

Inti dalam masalah penelitian ini ialah bagaimana penerapan metode algoritma *Naïve Bayes* pada transaksi-transaksi penjualan untuk menemukan solusi yang digunakan untuk mengoptimalkan skema jual beli di bidang kuliner. Hal ini dikarenakan butuhnya skema pemasaran yang lebih baik untuk menambahkan omset penjualan restoran. Persoalan yang dialami oleh restoran sejak dahulu ialah masih menggunakan cara biasa, yaitu hanya menggunakan cara mulut ke mulut saja untuk pemasaran, serta kurangnya pengetahuan tentang teknik pengolahan data penjualan yang baik dalam mengatasi solusi terbaik dari permasalahan tersebut.

Penerapan Data Mining Pada

## 2. TINJAUAN PUSTAKA

### 2.1. Data Mining

Data mining adalah sebuah jaringan dalam data yang ada dengan cara yang dapat dipahami sehingga menjadi informasi yang penting [1]. Teknik data mining dalam metode algoritma Naïve Bayes yang merupakan sebuah klasifikasi probabilitas sederhana. Basis dari dasar teori Bayes adalah mengatasi persoalan yang mempunyai sifat dugaan yaitu membuat suatu klasifikasi untuk menyisihkan objek. Data mining disebut juga sebagai Knowledge Discovery in Database (KDD), yaitu aktivitas yang meliputi pengumpulan, pembersihan data, penyesuaian data, penggalian data, dan tahap mempresentasikan data tersebut dalam set data berukuran besar [2][3].

### 2.2. Definisi Klasifikasi

Klasifikasi adalah salah satu metode yang digunakan dalam data mining. Klasifikasi dilakukan dengan memprediksi suatu class yang belum diketahui. Klasifikasi dapat dideskripsikan sebagai metode untuk membuktikan sebuah objek data sebagai salah satu jenis yang telah dideskripsikan sebelumnya [3].

### 2.3. Algoritma Naïve Bayes

Dalam teorema Bayes, probabilitas dapat dinyatakan sebagai berikut :

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

Dimana X adalah bukti, H adalah hipotesis, P(H|X) adalah probabilitas bahwa hipotesis H benar untuk bukti X atau dengan kata lain P(H|X) merupakan probabilitas posterior H dengan syarat X. P(X|H) adalah probabilitas bahwa hipotesis X benar untuk bukti H atau dengan kata lain P(X|H) merupakan probabilitas posterior X dengan syarat H. P(H) adalah probabilitas prior hipotesis H, dan P(X) adalah probabilitas prior bukti X.

### 2.4. Knowledge Discovery in Database (KDD)

Definisi Knowledge Discovery in Database (KDD) adalah metode dan cara mendapatkan sebuah informasi melalui basis data yang telah tersedia. Berikut tahap penyelesaiannya, yaitu :

#### 1. Data Selection

Data selection adalah pengambilan data dari sebuah dataset yang nantinya akan diolah.

#### 2. Data Preprocessing

Data preprocessing adalah tahap untuk melakukan sebuah proses awal dalam pengolahan data. Pada tahap ini data yang akan diolah bertujuan untuk menghindari dari data yang mengganggu (*noise*) atau data yang tidak konsisten.

#### 3. Transformation

Transformation adalah tahap membuat data yang akan diproses menjadi sesuai dengan model ataupun algoritma yang ingin digunakan dalam tahap pemrosesan data.

#### 4. Data Mining

Proses pencarian dan penggalian pengetahuan hingga diperoleh suatu model yang dapat digunakan untuk menjadi sebuah informasi penting dan berguna.

#### 5. Evaluation

Merepresentasikan hasil model yang telah diperoleh serta menguji akurasi dan kesesuaian terhadap data-data yang bersinggungan.

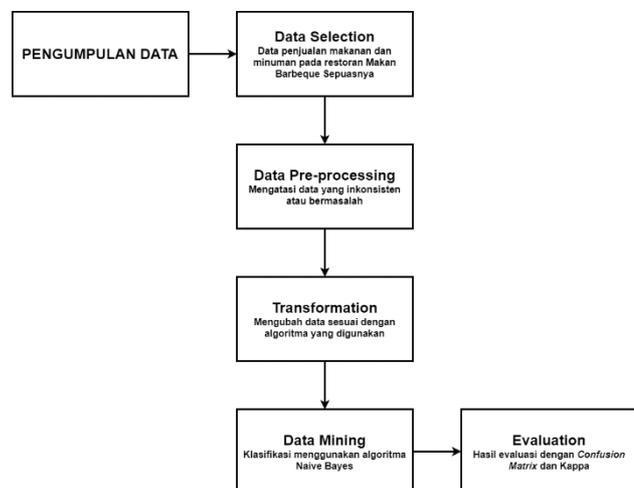
## 3. METODOLOGI

### 3.1. Objek Penelitian

Objek penelitian yang akan diteliti yaitu mengenai pengelompokan penjualan makanan dan minuman pada restoran Makan Barbeque Sepuasnya. Pengelompokan ini dilakukan pada pelanggan yang telah melakukan transaksi.

### 3.2. Metodologi Penelitian

Metodologi penelitian yang digunakan adalah Knowledge Discovery in Database (KDD). Tahapan pada metodologi KDD yaitu data selection, data preprocessing, transformation, data mining, evaluation [3].



Gambar 2 Alur Penelitian

### 3.3. Rancangan Penelitian

Untuk menyelesaikan permasalahan terkait dengan penggunaan algoritma Naïve Bayes, maka digunakanlah langkah-langkah Knowledge Discovery in Database (KDD) sebagai berikut [4][5]:

#### 1. Pengumpulan Data

Pengumpulan data pada penelitian ini adalah wawancara secara langsung kepada narasumber selaku pemilik restoran Makan Barbeque Sepuasnya. Peneliti dianjurkan mewawancarai orang yang akrab atau mengenai suatu topik pembahasan.

#### 2. Data Selection

Tahapan ini merupakan tahapan persiapan dalam pemilihan data. Data yang telah didapatkan dari wawancara akan dilakukan proses pemilihan atribut untuk selanjutnya digunakan untuk preprocessing.

#### 3. Data Preprocessing

Data Preprocessing bertujuan untuk mengubah data mentah menjadi data yang berkualitas sehingga data layak untuk diolah pada tahapan selanjutnya. Tahapan ini dilakukan pada data mentah untuk menghilangkan data yang bermasalah atau

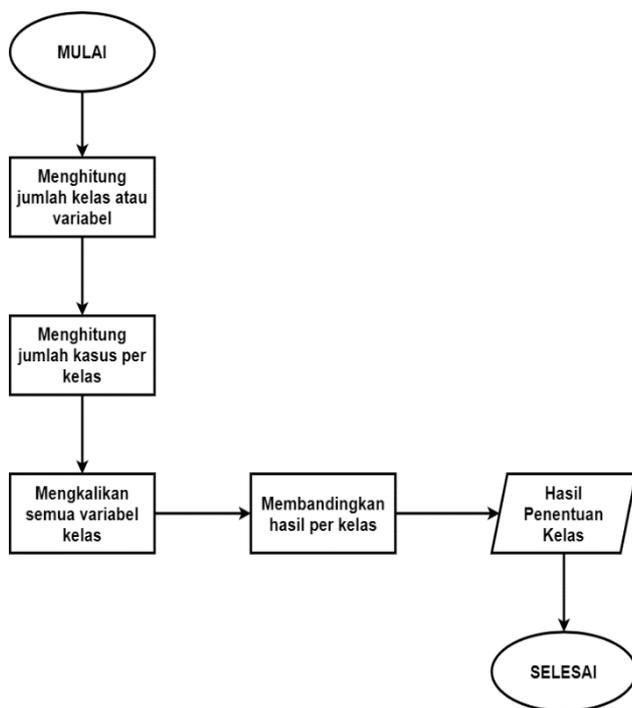
inkonsisten. Data yang bermasalah seperti data yang noise atau yang mengandung *error* [5].

4. Transformation

Transformation adalah tahapan di mana data diubah ke dalam bentuk yang sesuai dengan kebutuhan. Teknik transformation yang digunakan adalah *discretization*. Teknik bertujuan untuk mengubah atribut numerik menjadi atribut kategorikal, sehingga menjadi beberapa tingkatan kategori [5].

5. Data Mining

Pada tahapan ini data sudah siap untuk proses klasifikasi. Klasifikasi penjualan makanan dan minuman pada penelitian ini menggunakan aplikasi *RapidMiner*. Metode yang digunakan adalah metode algoritma *Naïve Bayes*. Metode algoritma *Naïve Bayes* digunakan untuk memprediksi peluang dimasa depan berdasarkan pengalaman dimasa sebelumnya, karena pengelompokkan penjualan makanan dan minuman ini dapat dilihat dari transaksi yang telah dilakukan sebelumnya untuk dijadikan prediksi atau peluang dimasa depan. Alur kerja dari metode algoritma *Naïve Bayes*, yaitu pertama membaca data training yang sudah siap digunakan, kedua menghitung jumlah kelas atau tabel, ketiga menghitung jumlah kasus per kelas, keempat mengkalikan semua variabel kelas dan bandingkan hasil per kelas. Hasil penentuan kelas dari probabilitas yang paling tinggi [6].



Gambar 3 Alur Algoritma *Naïve Bayes*

6. Evaluation

Pada tahap ini dilakukan evaluasi untuk mengukur validitas model klasifikasi dengan menggunakan *cross validation* dan mengevaluasi hasil penelitian dari perhitungan akurasi yang didapatkan dari *confusion matrix* dan nilai *Kappa* [7].

4. HASIL DAN PEMBAHASAN

4.1. Pengumpulan Data

Tahapan pengumpulan data ini dilakukan untuk mengumpulkan data penjualan yang diperoleh melalui wawancara kepada narasumber selaku pemilik restoran Makan Barbeque Sepuasnya

Tabel 1 Data Penjualan Makanan dan Minuman

Bulan	Makanan	Minuman
Februari	11373000	3791000
Maret	12750000	4250000
April	2550000	850000
Mei	3060000	1020000
Juni	15300000	5100000
Juli	17850000	5950000
Agustus	25500000	8500000
September	38250000	12750000
Oktober	22950000	7650000
November	22950000	7650000
Desember	22236xxx	7412xxx

Setelah data didapatkan, selanjutnya mendeskripsikan data tersebut yang bertujuan untuk memahami informasi lebih lanjut mengenai atribut dari data yang akan diolah dalam pengklasifikasian penjualan pada restoran Makan Barbeque Sepuasnya. Berikut adalah deskripsi atribut data yang telah didapatkan sebelumnya.

Tabel 2 Deskripsi Atribut Data

Atribut	Jenis	Keterangan
Makanan	Numeric	Nominal Penjualan Makanan
Minuman	Numeric	Nominal Penjualan Minuman

4.2. Data Selection

Data penjualan makanan dan minuman yang diambil melalui wawancara ini selajutnya memasuki tahap data selection. Tahapan data selection ini bertujuan memilih atribut yang dianggap sebagai atribut yang berpengaruh terhadap klasifikasi penjualan makanan dan minuman.

Tabel 3 Atribut Data Terpilih

Atribut
Makanan
Minuman

Objek penelitian yang akan diteliti yaitu mengenai pengelompokkan penjualan makanan dan minuman pada restoran Makan Barbeque Sepuasnya. Pengelompokkan ini dilakukan pada pelanggan yang telah melakukan transaksi.

### 4.3. Data Preprocessing

Pada data yang sudah diseleksi ternyata terdapat record ataupun data yang tidak sempurna. Pada tabel dibawah ini terdapat data yang berbeda tipe data dan data tersebut tidak valid atau salah dalam penulisan [7][8].

Tabel 4 Data Yang Tidak Relevan

Bulan	Makanan	Minuman
Februari	11373000	3791000
Maret	12750000	4250000
April	2550000	850000
Mei	3060000	1020000
Juni	15300000	5100000
Juli	17850000	5950000
Agustus	25500000	8500000
September	38250000	12750000
Oktober	22950000	7650000
November	22950000	7650000
Desember	22236xxx	7412xxx

Penghapusan missing value maupun data yang tidak relevan dilakukan agar tidak berpengaruh pada proses data mining sehingga menjadi data yang sempurna.

Tabel 5 Hasil Cleaning Data

Bulan	Makanan	Minuman
Februari	11373000	3791000
Maret	12750000	4250000
April	2550000	850000
Mei	3060000	1020000
Juni	15300000	5100000
Juli	17850000	5950000
Agustus	25500000	8500000
September	38250000	12750000
Oktober	22950000	7650000
November	22950000	7650000
Desember	22236000	7412000

Data yang berjenis numerik seperti makanan dan minuman harus dilakukan inisialisasi data terlebih dahulu ke dalam bentuk nominal makanan dan nominal minuman. Untuk melakukan inisialisasi dapat dilakukan dengan sebagai berikut :

1. Makanan yang terjual dengan nominal lebih dari 15.000.000 diberi inisial pada atribut Keterangan Makanan “Untung”.

2. Makanan yang terjual dengan nominal kurang dari atau sama dengan 15.000.000 diberi inisial pada atribut Keterangan Makanan “Rugi”.
3. Minuman yang terjual dengan nominal lebih dari 5.000.000 diberi inisial pada atribut Keterangan Minuman “Untung”.
4. Minuman yang terjual dengan nominal kurang dari atau sama dengan 5.000.000 diberi inisial pada atribut Keterangan Minuman “Rugi”
5. Berikut adalah dataset penjualan makanan dan minuman yang telah dilakukan inisialisasi yang terdapat pada gambar di bawah ini.

Tabel 6 Penginisialisasi Data Set

Bulan	Nominal Makanan	Keterangan Makanan	Nominal Minuman	Keterangan Minuman
Februari	11373000	Rugi	3791000	Rugi
Maret	12750000	Rugi	4250000	Rugi
April	2550000	Rugi	850000	Rugi
Mei	3060000	Rugi	1020000	Rugi
Juni	15300000	Untung	5100000	Untung
Juli	17850000	Untung	5950000	Untung
Agustus	25500000	Untung	8500000	Untung
September	38250000	Untung	12750000	Untung
Oktober	22950000	Untung	7650000	Untung
November	22950000	Untung	7650000	Untung
Desember	22236000	Untung	7412000	Untung

### 4.4. Transformation

Transformation merupakan tahapan perubahan data kedalam bentuk yang sesuai dengan kebutuhan. Pada tahap ini dilakukan transformasi untuk mengubah tipe data numerik menjadi kategori untuk mempermudah proses klasifikasi. Proses transformasi ini dilakukan dengan dengan bantuan operator *Discretization by Binning*. Teknik *discretization* dilakukan untuk mengubah nilai atribut kontinu menjadi sejumlah interval yang terbatas dan berasosiasi dengan masing-masing interval dengan nilai numerik yang diskrit. menjadi beberapa kelompok yang diinginkan [3]. Perubahan ini dilakukan untuk mengkategorikan atribut numerik menjadi beberapa tingkatan kategori. Pada penelitian ini tidak ada tahap transformation, karena tidak ada data untuk menjadi kategori [8][9].

### 4.5. Data Mining

Data penjualan makanan dan minuman yang telah siap selanjutnya akan dilakukan pemodelan data *mining* dengan algoritma *Naïve Bayes* untuk mengklasifikasikan keuntungan atau kerugian yang akan datang. Data *mining* dengan menggunakan algoritma *Naïve Bayes* memiliki beberapa tahapan yang pertama menghitung jumlah kelas dari data training yang sudah siap digunakan, lalu menghitung jumlah kasus dari setiap kelas yang ada, selanjutnya mengkalikan semua variabel kelas lalu membandingkan hasil dari setiap kelas dan menghasilkan output dari hasil penentuan kelas yang dilihat dari probabilitas yang paling tinggi [9][10].

### 4.5.1. Menghitung Jumlah Kelas

Dari 22 *cell* (data) yang siap digunakan dengan kelas “Keterangan Makanan” sebanyak 11 *cell* (data) dan “Keterangan Minuman” sebanyak 11 *cell* (data). Dari jumlah *cell* (data) dari masing masing kelas dibagi dengan keseluruhan *cell* (data) maka akan mendapatkan probabilitas *prior*. Berikut perhitungan probabilitas *prior* berdasarkan persamaan:

$$P(\text{Keterangan Makanan, Untung}) = 7/22 = 0,318$$

$$P(\text{Keterangan Makanan, Rugi}) = 4/22 = 0,182$$

$$P(\text{Keterangan Minuman, Untung}) = 7/22 = 0,318$$

$$P(\text{Keterangan Minuman, Rugi}) = 4/22 = 0,182$$

Tabel 7 Perhitungan Probabilitas *Prior*

Kelas	Sub Kelas	Jumlah	Probabilitas
			Kelas P (C)
Keterangan	Untung	7	0,318
Makanan	Rugi	4	0,182
Keuntungan	Untung	7	0,318
Minuman	Rugi	4	0,182
<b>Total</b>		<b>22</b>	<b>1,000</b>

Keterangan :

Probabilitas prior menggunakan symbol P (Ci)

### 4.5.2. Menghitung Jumlah Kasus Dari Setiap Kelas

Setelah menghitung probabilitas *prior* selanjutnya menghitung probabilitas *posterior* atau jumlah kasus maupun kejadian dari setiap kelas. Untuk mencari nilai probabilitas *posterior* yaitu jumlah atribut dengan kelas “Keuntungan Makanan” dan kelas “Keuntungan Minuman” kemudian dibagi dengan jumlah kelas yang ada.

Tabel 8 Probabilitas Pada Bulan

Bulan	Jumlah Kejadian			
	Keterangan Makanan		Keterangan Minuman	
	Untung	Rugi	Untung	Rugi
Februari	0	1	0	1
Maret	0	1	0	1
April	0	1	0	1
Mei	0	1	0	1
Juni	1	0	1	0
Juli	1	0	1	0
Agustus	1	0	1	0
September	1	0	1	0
Oktober	1	0	1	0
November	1	0	1	0
Desember	1	0	1	0
<b>Total</b>	<b>11</b>	<b>11</b>	<b>11</b>	<b>11</b>

Bulan	Probabilitas P (F C)			
	Keterangan Makanan		Keterangan Minuman	
	Untung	Rugi	Untung	Rugi
Februari	0	0,09	0	0,09
Maret	0	0,09	0	0,09
April	0	0,09	0	0,09
Mei	0	0,09	0	0,09
Juni	0,09	0	0,09	0
Juli	0,09	0	0,09	0
Agustus	0,09	0	0,09	0
September	0,09	0	0,09	0
Oktober	0,09	0	0,09	0
November	0,09	0	0,09	0
Desember	0,09	0	0,09	0
<b>Total</b>	<b>1,000</b>	<b>1,000</b>	<b>1,000</b>	<b>1,000</b>

### 4.5.3. Mengalikan Semua Variabel Kelas

Perhitungan nilai probabilitas *prior* dan probabilitas *posterior* yang telah dilakukan akan digunakan sebagai model yang akan digunakan sebagai acuan untuk menentukan data *testing*. Di bawah ini merupakan contoh data *testing* yang akan di hitung probabilitasnya.

1. Untuk semua atribut kelas Keterangan Makanan = “Untung”  
 $P(X | \text{Keterangan Makanan} = \text{“Untung”}) = 7/11 = 0,636$
2. Untuk semua atribut kelas Keterangan Makanan = “Rugi”  
 $P(X | \text{Keterangan Makanan} = \text{“Rugi”}) = 4/11 = 0,363$
3. Untuk semua atribut kelas Keterangan Minuman = “Untung”  
 $P(X | \text{Keterangan Minuman} = \text{“Untung”}) = 7/11 = 0,636$
4. Untuk semua atribut kelas Keterangan Minuman = “Rugi”  
 $P(X | \text{Keterangan Minuman} = \text{“Rugi”}) = 4/11 = 0,363$
5. Perkalian probabilitas prior dengan semua atribut Keterangan Makanan = “Untung”  
 $P(C_i | \text{Keterangan Makanan} = \text{“Untung”}) \times P(X | \text{Keterangan Makanan} = \text{“Untung”}) = 0,318 \times 0,636 = 0,202$
6. Perkalian probabilitas prior dengan semua atribut Keterangan Makanan = “Rugi”  
 $P(C_i | \text{Keterangan Makanan} = \text{“Rugi”}) \times P(X | \text{Keterangan Makanan} = \text{“Rugi”}) = 0,182 \times 0,363 = 0,066$
7. Perkalian probabilitas prior dengan semua atribut Keterangan Minuman = “Untung”  
 $P(C_i | \text{Keterangan Minuman} = \text{“Untung”}) \times P(X | \text{Keterangan Minuman} = \text{“Untung”}) = 0,318 \times 0,636 = 0,202$
8. Perkalian probabilitas prior dengan semua atribut Keterangan Minuman = “Rugi”  
 $P(C_i | \text{Keterangan Minuman} = \text{“Rugi”}) \times P(X | \text{Keterangan Minuman} = \text{“Rugi”}) = 0,182 \times 0,363 = 0,066$

#### 4.5.4. Membandingkan Hasil Dari Setiap Kelas

Didapatkan hasil perhitungan probabilitas data testing dari setiap kelas yang digunakan pada tabel dibawah ini.

Tabel 9 Hasil Perhitungan Probabilitas Data Testing

Kelas	Probabilitas
P (Keterangan Makanan   Untung)	0,202
P (Keterangan Makanan   Rugi)	0,066
P (Keterangan Minuman   Untung)	0,202
P (Keterangan Minuman   Rugi)	0,066

Dari data *testing* yang sudah diuji menghasilkan kelas P (Keterangan Makanan | Untung) dan P (Keterangan Minuman | Untung) memiliki nilai probabilitas tertinggi diantara kelas lainnya, sehingga dapat diambil kesimpulan bahwa makanan dan minuman memiliki probabilitas lebih tinggi.

#### 4.6. Evaluation

Tahapan evaluasi dilakukan untuk menilai sejauh mana data yang telah dibuat sehingga dapat memenuhi tujuan klasifikasi. Hasil dari pengujian metode *Naïve Bayes* berdasarkan skenario yang telah dibuat dengan *cross validation* mendapatkan nilai *confusion matrix* dan nilai *kappa*.

Tabel 10 Hasil Pengujian Berdasarkan *Accuracy*

Skenario	2 fold	3 fold	4 fold	5 fold	6 fold	7 fold	8 fold	9 fold	10 fold
Accuracy	85,92	88,73	88,03	87,32	88,38	87,68	86,9	88,03	88,3

Tabel 11 Hasil Pengujian Berdasarkan *Precision*

Skenario	2 fold	3 fold	4 fold	5 fold	6 fold	7 fold	8 fold	9 fold	10 fold
Precision	45,56	64,62	56,37	54,43	58,38	47,78	56,8	58,03	66,37

Tabel 12 Hasil Pengujian Berdasarkan *Recall*

Skenario	2 fold	3 fold	4 fold	5 fold	6 fold	7 fold	8 fold	9 fold	10 fold
Recall	45,12	45,72	43,13	43,52	41,36	42,78	44,9	44,83	48,7

Tabel 13 Hasil Pengujian Berdasarkan Nilai *Kappa*

Skenario	2 fold	3 fold	4 fold	5 fold	6 fold	7 fold	8 fold	9 fold	10 fold
Kappa	0,392	0,425	0,453	0,21	0,34	0,444	0,42	0,411	0,431

### 5. KESIMPULAN DAN SARAN

Hasil dari penelitian yang telah dilakukan, maka dapat diambil kesimpulan sebagai berikut:

1. Penelitian ini dilakukan untuk mengklasifikasikan penjualan makanan dan minum pada restoran Makan Barbeque Sepuasnya menggunakan algoritma *Naïve Bayes* dengan metodologi KDD (*Knowledge Discovery in Database*). Dari hasil perhitungan klasifikasi menggunakan algoritma *Naïve Bayes* adalah keterangan makanan untung dan keterangan minuman untung yang lebih besar probabilitasnya.
2. Hasil pengujian klasifikasi algoritma *Naïve Bayes* dari sembilan skenario pengujian yang telah dibuat dengan *cross validation*, menghasilkan performa terbaik pada skenario pengujian dengan menggunakan 3 *fold* yang menghasilkan performa terbaik dengan nilai *accuracy* sebesar 88,73%,

*precision* sebesar 64,42%, *recall* sebesar 45,41% dan dengan nilai *kappa* yang diperoleh sebesar 0,451 yang termasuk kedalam kategori cukup. Berdasarkan hasil tersebut maka, model yang dihasilkan oleh algoritma *Naïve Bayes* ini konsisten.

Dari penelitian yang telah dilakukan, terdapat saran yang perlu disampaikan pada penelitian ini untuk referensi penelitian yang akan datang :

1. Perbanyak dataset dan juga atribut yang lebih berpengaruh untuk klasifikasi penjualan makanan dan minuman agar akurasi algoritma *Naïve Bayes* semakin meningkat.
2. Pada penelitian ini perlu ditambahkan teknik-teknik lainnya untuk mendapatkan performa klasifikasi yang lebih sempurna untuk menangani data yang tidak seimbang.
3. Perlu ditambahkan skenario pengujian lainnya dalam membangun sebuah model untuk mendapatkan hasil yang lebih baik

### DAFTAR PUSTAKA

- [1] Nurdiawan, Odi., Noval Salim. 2018. “Penerapan Data Mining Pada Penjualan Barang Menggunakan Metode *Naïve Bayes Classifier* Untuk Optimasi Strategi Pemasaran.” *Jurnal Teknologi Informasi dan Komunikasi*.
- [2] Ridwan, Mujib., Hadi Suyono., M. Sarosa. 2013. “Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma *Naïve Bayes Classifier*.” *Jurnal EECIS Vol. 7 No.1*.
- [3] Mardi, Yuli. 2017. “Data Mining: Klasifikasi menggunakan Algoritma C4.5.” *Jurnal Edik Informatika V2.i2 (213-219)*
- [4] Harwati., Hanna Miratama. 2016. “Prediksi Prestasi Mahasiswa Dengan Jalur Siswa Berprestasi (PSB) Menggunakan Metode *Naïve Bayes*.”
- [5] Effendi, M.R. 2015. “Akurasi Data Mining Untuk Menghasilkan Pola Kelulusan Mahasiswa Dengan Metode *Naïve Bayes*.” 101-106.
- [6] Rizqiyani, Vidya., Anggraini Mulwinda., Riana Defi Mahadji Putri. 2017. “Klasifikasi Judul Buku dengan Algoritma *Naïve Bayes* dan Pencarian Buku Pada Perpustakaan Jurusan Teknik Elektro.” *Jurnal Teknik Elektro Vol. 9 No. 2*.
- [7] Anggraini, Mathia., Rizki Ayuning Tyas., Ismi Ana Sulasyah., Qurrotul Aini. 2020. “Implementasi Algoritma *Naïve Bayes* Dalam Penentuan Rating Buku.” *Jurnal Sistem Informasi Vol. 9 No. 3*.
- [8] Watratan, Alvina Felicia., Arwini Puspita B. ., Dikwan Moeis. 2020. “Implementasi Algoritma *Naïve Bayes* Untuk Memprediksi Tingkat Penyebaran Covid-19 Di Indonesia.” *Journal Of Applied Computer Science And Technology (Jacost) Vol. 1 No. 1*.
- [9] Saleh, Alfa. 2015. “Implementasi Metode Klasifikasi *Naïve Bayes* Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga.” *Citec Journal Vol. 2 No. 3*.
- [10] Maryah., M. Fadel Asikin., Daisy Kurniawaty., Selly Kurnia Sari., Imam Cholissodin. 2016. “Implementasi Metode *Naïve Bayes Classifier* Untuk Seleksi Asisten Praktikum Pada Simulasi Hadoop Multinode Cluster.” *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK) Vol. 3 No. 4*.

## BIODATA PENULIS



**Fauzan Alghifari**

Mahasiswa Teknik Informatika, Jalan H.S.  
Ronggowaluyo Kel. Puseurjaya Kec.  
Telukjambe Timur Kab. Karawang Prov.  
Jawa Barat 41361.



**Didi Juardi , S.T., M.Kom.**

Dosen Teknik Informatika, Jalan H.S.  
Ronggowaluyo Kel. Puseurjaya Kec.  
Telukjambe Timur Kab. Karawang Prov.  
Jawa Barat 41361.