

# MEMPREDIKSI TINGKAT KEPUASAN PELANGGAN TERHADAP KUALITAS SUATU TOKO MENGGUNAKAN METODE NAÏVE BAYES

**Rizki Delaga Prasetya<sup>1</sup>, Raihan Iqbal Pasya<sup>2</sup>, Rafi Awang Bagaskoro<sup>3</sup>**

<sup>1,2,3</sup>Institut Teknologi Telkom Purwokerto,

Jl. DI Panjaitan No.128, Kabupaten Banyumas, Jawa Tengah 53147

<sup>1</sup>19104074@ittelkom-pwt.ac.id, <sup>2</sup>19104030@ittelkom-pwt.ac.id, <sup>3</sup>19104018@ittelkom-pwt.ac.id

## ABSTRAK

*Naïve Bayes* merupakan metode klasifikasi probabilitas untuk memprediksi peluang yang akan terjadi berdasarkan pengalaman di masa lalu yang sudah terjadi. Penelitian ini bertujuan untuk memprediksi tingkat kepuasan pelanggan terhadap kualitas suatu toko berdasarkan beberapa komponen seperti Harga, Keamanan, Lokasi, Fasilitas, *Service*, dan Kelengkapan Barang menggunakan metode *Naïve Bayes*. Sampel yang digunakan pada penelitian ini diambil sebesar 30 data yang diperoleh dari wawancara kepada beberapa teman-teman dan pendapat dari penulis sendiri terhadap beberapa toko yang pernah dikunjungi. Pada penelitian ini, penulis menggunakan metode *Naïve Bayes* dikarenakan metode ini tidak membutuhkan jumlah data yang banyak untuk bisa melakukan prediksi, dan *Naïve Bayes* juga dikenal memiliki tingkat akurasi bagus. Pengujian dilakukan dengan cara manual menggunakan rumus Bayes yang dipopulerkan oleh Paul Graham dan yang kedua menggunakan program Artificial Intelligence(AI) yang menerapkan metode Gaussian. Hasil dari penelitian menunjukkan bahwa prediksi menggunakan metode *Naïve Bayes* memiliki tingkat akurasi yang bagus, itu dibuktikan dengan hasil akurasi prediksi mencapai 80%, dihitung menggunakan perhitungan manual maupun menggunakan Program *Machine Learning*.

**Kata kunci-** *Customer Satisfaction, Machine Learning, Naïve Bayes, Probability*

## I. PENDAHULUAN

Toko merupakan sebuah tempat yang di dalamnya terjadi suatu kegiatan perdagangan jual beli barang maupun jasa. Kualitas suatu toko bisa dipengaruhi oleh banyak komponen seperti Harga, Keamanan, Lokasi, Fasilitas, *Service*, dan Kelengkapan Barang. Peran pelanggan sebagai pembeli sangat menentukan reputasi sebuah toko. Reputasi sebuah toko bisa dilihat dari tingkat kepuasan pelanggan terhadap kualitas suatu toko. Kepuasan pelanggan ditentukan oleh persepsi pelanggan atas kinerja produk atau jasa dalam memenuhi harapannya. Kepuasan pelanggan akan tercapai apabila setelah pelanggan menggunakan suatu produk atau jasa, apa yang diinginkan dan dibutuhkan pelanggan dapat terpenuhi bahkan melebihi harapannya [1].

Pada penelitian kali ini penulis mencoba untuk memprediksi tingkat kepuasan pelanggan terhadap kualitas suatu toko dengan menerapkan metode *classification Naïve Bayes*.

Istilah dari kata Prediksi sendiri bisa menimbulkan berbagai macam persepsi. Prediksi atau dalam hal ini juga disebut perkiraan (*estimate*) pada dasarnya merupakan dugaan atau perkiraan mengenai suatu peristiwa di waktu yang akan datang [2]. *Naïve Bayes* adalah sebuah metode klasifikasi yang menggunakan probabilitas dan *statistic*. Dengan kata lain dalam *Naïve Bayes* kita memprediksi sebuah peluang yang akan terjadi berdasarkan pengalaman di masa lalu yang sudah terjadi.

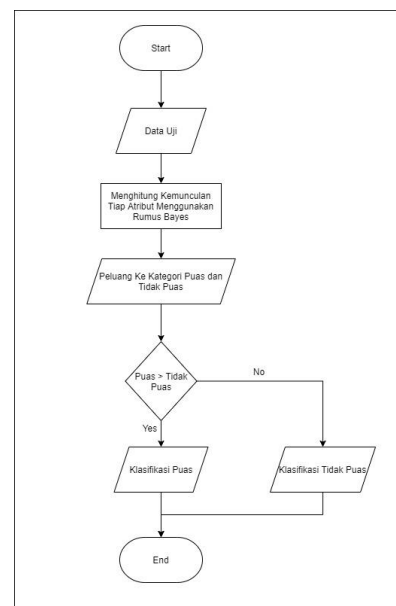
Kegunaan *Naïve Bayes* pada penelitian kali ini sebagai metode *machine learning* yang menggunakan probabilitas untuk memprediksi tingkat kepuasan pelanggan terhadap sebuah toko. *Naïve Bayes* merupakan metode paling sederhana dari pengklasifikasian probabilitas, memiliki tingkat akurasi yang sangat tinggi ketika diaplikasikan pada *database* dengan *big data* [3].

Penelitian ini dilakukan untuk mencari tahu kualitas suatu toko dari tingkat kepuasan pelanggan terhadap toko. Analisis ini menggunakan *Naïve Bayes* sebagai metode untuk menghitung tingkat kepuasan pelanggan terhadap suatu toko.

Dari uraian seelumnya perlu melakukan penelitian tentang prediksi tingkat kepuasan pelanggan terhadap kualitas suatu toko menggunakan metode *Naïve Bayes*. Yang bertujuan untuk dapat meningkatkan kualitas dari toko-toko yang ada di Indonesia.

## II. METODOLOGI PENELITIAN

### 2.1. Flowchart



Gambar 1. *Flowchart Naïve Bayes Untuk Menghitung Tingkat Kepuasan*

Alur dari pengklasifikasian menggunakan metode *Naïve Bayes* yaitu :

- 1) Melakukan uji data terhadap data yang ada.
- 2) Menghitung kemunculan tiap atribut sesuai dengan nilai atribut yang dicari.
- 3) Mendefinisikan nilai perhitungan kedalam kategori PUAS atau TIDAK PUAS menggunakan rumus *Bayes*.
- 4) Jika dari hasil perhitungan, nilai PUAS pada tingkat kepuasan lebih tinggi dari pada nilai TIDAK PUAS berdasarkan data-data yang sudah ada, maka tingkat kepuasan dari data yang dicari termasuk dalam kategori PUAS, begitu juga sebaliknya.

## 2.2. Data

*Sample* data didapatkan dari hasil wawancara terhadap pembeli mengenai beberapa toko yang pernah dia kunjungi, dengan memberikan penilaian terhadap toko tersebut sesuai dengan atribut yang telah ditetapkan. Enam indikator yang digunakan sebagai atribut untuk mengukur tingkat kepuasan pelanggan terhadap suatu toko yaitu:

### 1) Harga

Barang yang berkualitas dengan harga bersaing di pasaran merupakan kunci utama dalam memenangkan persaingan, yang pada akhirnya toko akan mendapatkan nilai kepuasan yang lebih tinggi dari pembelinya [4].

### 2) Keamanan

Keamanan akan memberikan kepercayaan dan kenyamanan bagi pembeli yang datang. Pelanggan akan senang hati untuk terus berbelanja di toko yang memiliki keamanan yang bagus, sehingga toko yang memiliki keamanan yang bagus akan banyak disukai dengan pelanggannya.

### 3) Lokasi

Lokasi yang strategis akan mempermudah pembeli dalam menemukan toko yang dicarinya. Ciri-ciri lokasi yang strategis antara lain yaitu lokasi mudah terlihat, tempat mudah dijangkau, letak toko yang ideal, dekat dengan fasilitas umum dan lain-lain.

### 4) Fasilitas

Fasilitas seperti parkir gratis, keranjang barang, *CCTV*, toilet, dan kursi akan membuat pembeli merasa terbantu dengan adanya fasilitas tersebut.

### 5) Service

Pelayanan terhadap pembeli merupakan salah satu faktor yang menjadikan suatu toko banyak memiliki pembeli setia. Kemudahan pembayaran, jumlah karyawan, keramahan, respon dan dari komunikasi karyawan akan sangat disukai oleh pembeli.

### 6) Kelengkapan Barang

Toko yang menyediakan barang yang lengkap akan banyak didatangi pembeli, karena barang-barang yang dibutuhkan pembeli ada di toko tersebut sehingga

pembeli tidak perlu mencari ke toko lain untuk membeli barang yang dibutuhkan.

## 2.3. Metode yang Digunakan

Pada penelitian ini akan mencari tahu kualitas suatu toko dari tingkat kepuasan pelanggan terhadap toko menggunakan metode *Classifier Naïve Bayes*. *Naïve Bayes* adalah sebuah metode klasifikasi yang menggunakan probabilitas dan *statistic*. Dengan kata lain dalam *Naïve Bayes* kita memprediksi sebuah peluang yang akan terjadi berdasarkan pengalaman di masa lalu yang sudah terjadi. Kegunaan *Naïve Bayes* pada penelitian kali ini sebagai metode *machine learning* yang menggunakan probabilitas untuk memprediksi tingkat kepuasan pelanggan terhadap sebuah toko.

*Naïve Bayes* merupakan metode paling sederhana dari pengklasifikasian probabilitas, memiliki tingkat akurasi yang sangat tinggi ketika diaplikasikan pada *database* dengan *big data* [3]. *Naïve Bayes* merupakan metode pengklasifikasian statistik yang digunakan untuk memprediksi probabilitas keanggotaan dari suatu *class*. Dalam beberapa kasus, algoritma *Naïve Bayes* terbukti memiliki tingkat akurasi dan kecepatan yang tinggi saat diaplikasikan ke dalam *database* dengan data yang besar [5].

Kelebihan dari *Naïve Bayes* antara lain yaitu :

- 1) Perhitungan yang cepat dan efisien sehingga tidak membutuhkan proses dan waktu yang lama untuk memprediksi tingkat kepuasan pelanggan.
- 2) Tidak memerlukan jumlah data yang banyak, hal ini membuat proses pengumpulan data menjadi lebih efisien karena tidak membutuhkan jumlah data yang terlalu banyak.
- 3) Simple dan Efektif, karena *code* program yang sederhana tetapi dapat digunakan untuk masalah *biner* ataupun *multiclass*.

## 2.4. Pengujian

Berdasarkan metode – metode yang sudah dijelaskan sebelumnya, akan melakukan pengujian dengan dua cara yaitu dengan perhitungan secara manual menggunakan rumus *Bayes* yang dipopulerkan oleh Paul Graham, dan yang kedua menggunakan program *AI (Artificial Intelligence)* yang menerapkan metode Gaussian dalam memprediksi dan menghitung studi kasus *Naïve Bayes*.

### 1) Perhitungan Manual menggunakan Rumus *Bayes*

Pada penelitian kali ini kita akan mencoba untuk memprediksi tingkat kepuasan seseorang terhadap atribut Harga, Keamanan, Lokasi, Fasilitas, *Service* dan Kelengkapan Barang dari sebuah toko. Berikut data tabel yang sudah didapatkan dari beberapa pelanggan mengenai tingkat kepuasan seseorang terhadap sebuah toko.

Tingkat Kepuasan	Harga	Keamanan	Lokasi	Fasilitas	Service	Kelengkapan Barang
Puas	Mahal	Baik	Strategis	Baik	Cukup	Sangat Baik
Tidak Puas	Mahal	Kurang	Tidak Strategis	Cukup	Baik	Cukup
Puas	Murah	Baik	Strategis	Cukup	Baik	Baik
Tidak Puas	Mahal	Baik	Strategis	Baik	Kurang	Cukup
Tidak Puas	Mahal	Kurang	Tidak Strategis	Buruk	Kurang	Kurang
Puas	Murah	Baik	Strategis	Baik	Sangat Baik	Baik
Puas	Murah	Cukup	Tidak Strategis	Baik	Cukup	Baik
Puas	Mahal	Cukup	Strategis	Baik	Baik	Baik
Tidak Puas	Mahal	Cukup	Strategis	Kurang	Buruk	Cukup
Tidak Puas	Mahal	Kurang	Tidak Strategis	Cukup	Kurang	Cukup
Puas	Murah	Cukup	Strategis	Baik	Sangat Baik	Cukup
Tidak Puas	Murah	Kurang	Tidak Strategis	Cukup	Cukup	Baik
Puas	Murah	Baik	Strategis	Cukup	Baik	Baik
Tidak Puas	Mahal	Baik	Strategis	Baik	Kurang	Cukup
Tidak Puas	Murah	Kurang	Tidak Strategis	Kurang	Buruk	Kurang
Puas	Murah	Baik	Strategis	Sangat Baik	Baik	Baik
Puas	Murah	Baik	Tidak Strategis	Cukup	Baik	Baik
Puas	Mahal	Cukup	Strategis	Baik	Baik	Baik
Tidak Puas	Mahal	Cukup	Strategis	Kurang	Buruk	Cukup
Tidak Puas	Murah	Buruk	Tidak Strategis	Cukup	Kurang	Cukup
Tidak Puas	Murah	Sangat Baik	Strategis	Baik	Baik	Baik
Tidak Puas	Murah	Cukup	Strategis	Cukup	Cukup	Cukup
Tidak Puas	Murah	Cukup	Tidak Strategis	Baik	Baik	Kurang
Tidak Puas	Murah	Baik	Strategis	Cukup	Baik	Baik
Tidak Puas	Murah	Cukup	Strategis	Cukup	Cukup	Baik
Puas	Murah	Cukup	Strategis	Kurang	Cukup	Baik
Puas	Mahal	Kurang	Tidak Strategis	Buruk	Kurang	Buruk
Tidak Puas	Murah	Kurang	Strategis	Cukup	Kurang	Buruk
Puas	Murah	Kurang	Tidak Strategis	Cukup	Cukup	Baik
Puas	Murah	Cukup	Tidak Strategis	Cukup	Baik	Baik

Gambar 2. Dataset

Selanjutnya, kita akan menghitung probabilitas kemunculan dari setiap atribut/kolom pada tabel tersebut seperti berikut :

- a. Probabilitas kemunculan setiap nilai untuk atribut Harga

Harga	Jumlah Kejadian Dipilih		Probabilitas	
	Puas	Tidak Puas	Puas	Tidak Puas
Murah	10	9	10/14	9/16
Mahal	4	7	4/14	7/16
<b>Total</b>	<b>14</b>	<b>16</b>	<b>1</b>	<b>1</b>

Gambar 3. Tabel Probabilitas untuk atribut Harga

- b. Probabilitas kemunculan setiap nilai untuk atribut Keamanan

Keamanan	Jumlah Kejadian Dipilih		Probabilitas	
	Puas	Tidak Puas	Puas	Tidak Puas
Sangat Baik	0	1	0	1/16
Baik	6	3	6/14	3/16
Cukup	6	5	6/14	5/16
Kurang	2	6	2/14	6/16
Buruk	0	1	0	1/16
<b>Total</b>	<b>14</b>	<b>16</b>	<b>1</b>	<b>1</b>

Gambar 4. Tabel Probabilitas untuk atribut Keamanan

- c. Probabilitas kemunculan setiap nilai untuk atribut Lokasi

Lokasi	Jumlah Kejadian Dipilih		Probabilitas	
	Puas	Tidak Puas	Puas	Tidak Puas
Strategis	9	9	9/14	9/16
Tidak Strategis	5	7	3/14	9/16
<b>Total</b>	<b>14</b>	<b>16</b>	<b>1</b>	<b>1</b>

Gambar 5. Tabel Probabilitas untuk atribut Lokasi

- d. Probabilitas kemunculan setiap nilai untuk atribut Fasilitas

Fasilitas	Jumlah Kejadian Dipilih		Probabilitas	
	Puas	Tidak Puas	Puas	Tidak Puas
Sangat Baik	1	0	1/14	0
Baik	6	4	6/14	4/16
Cukup	5	8	5/14	8/16
Kurang	1	3	1/14	3/16
Buruk	1	1	1/14	1/16
<b>Total</b>	<b>14</b>	<b>16</b>	<b>1</b>	<b>1</b>

Gambar 6. Tabel Probabilitas untuk atribut Fasilitas

e. Probabilitas kemunculan setiap nilai untuk atribut *Service*

Service	Jumlah Kejadian Dipilih		Probabilitas	
	Puas	Tidak Puas	Puas	Tidak Puas
Sangat Baik	2	0	2/14	0
Baik	7	4	7/14	4/16
Cukup	4	3	4/14	3/16
Kurang	1	6	1/14	6/16
Buruk	0	3	0	3/16
<b>Total</b>	<b>14</b>	<b>16</b>	<b>1</b>	<b>1</b>

Gambar 7. Tabel Probabilitas untuk atribut *Service*

f. Probabilitas kemunculan setiap nilai untuk atribut Kelengkapan Barang

Kelengkapan Barang	Jumlah Kejadian Dipilih		Probabilitas	
	Puas	Tidak Puas	Puas	Tidak Puas
Sangat Baik	1	0	1/14	0
Baik	11	4	11/14	4/16
Cukup	1	8	1/14	8/16
Kurang	0	3	0	3/16
Buruk	1	1	1/14	1/16
<b>Total</b>	<b>14</b>	<b>16</b>	<b>1</b>	<b>1</b>

Gambar 8. Tabel Probabilitas untuk atribut Kelengkapan Barang

g. Probabilitas kemunculan setiap nilai untuk atribut Tingkat Kepuasan

Tingkat Kepuasan	Jumlah Kejadian Dipilih		Probabilitas	
	Puas	Tidak Puas	Puas	Tidak Puas
<b>Total</b>	<b>14</b>	<b>16</b>	<b>14/30</b>	<b>16/30</b>

Gambar 9. Tabel Probabilitas untuk atribut Tingkat Kepuasan

Setelah mendapatkan nilai probabilitas setiap atribut, maka kita dapat menghitung tingkat kepuasan seseorang terhadap suatu toko dengan menggunakan rumus *Bayes* berikut:

$$P(A | B) = (P(B | A) * P(A)) / P(B)$$

Keterangan[3]:

- B : Data dengan kelas yang belum diketahui.
- A : Hipotesa data B merupakan suatu kelas spesifik.
- P(A) : Probabilitas hipotesa A (*prior probability* / probabilitas awal).
- P(B) : Probabilitas B.
- P(B|A) : Probabilitas hipotesa B berdasarkan kondisi A.
- P(A|B) : Probabilitas hipotesa A berdasarkan kondisi B (*Posterior probability* / probabilitas akhir).

Sekarang mari kita coba untuk menghitung probabilitas tingkat kepuasan seseorang dari suatu kasus, jika ia merasa Harga barang di suatu toko tersebut MURAH, Keamanan di toko tersebut CUKUP, Lokasi toko tersebut STRATEGIS, memberikan Fasilitas yang KURANG, memberikan *Service* yang CUKUP, dan memiliki Kelengkapan Barang yang BAIK. Maka dapat

kita hitung probabilitas puas dan probabilitas tidak seperti gambar 10 dan gambar 11.

$$\begin{aligned}
 \text{PUAS} &= P(\text{Puas}|\text{Harga} = \text{MURAH}) \cdot P(\text{Puas}|\text{Keamanan} = \text{CUKUP}) \cdot P(\text{Puas}|\text{Lokasi} = \\
 &\text{STRATEGIS}) \cdot P(\text{Puas}|\text{Fasilitas} = \text{KURANG}) \cdot P(\text{Puas}|\text{Service} = \text{CUKUP}) \cdot \\
 &P(\text{Puas}|\text{Kelengkapan Barang} = \text{BAIK}) \cdot P(\text{PUAS}) \\
 &= 10/14 * 6/14 * 9/14 * 1/14 * 4/14 * 11/14 * 14/30 \\
 &= 0.0014726
 \end{aligned}$$

Gambar 10. Perhitungan Nilai Puas

$$\begin{aligned}
 \text{TIDAK} &= (Tidak|\text{Harga} = \text{MURAH}) \cdot P(\text{Tidak}|\text{Keamanan} = \text{CUKUP}) \cdot P(\text{Tidak}|\text{Lokasi} = \\
 &\text{STRATEGIS}) \cdot P(\text{Tidak}|\text{Fasilitas} = \text{KURANG}) \cdot P(\text{Tidak}|\text{Service} = \text{CUKUP}) \cdot \\
 &P(\text{Tidak}|\text{Kelengkapan Barang} = \text{BAIK}) \cdot P(\text{TIDAK}) \\
 &= 9/16 * 5/16 * 9/16 * 3/16 * 3/16 * 4/16 * 16/30 \\
 &= 0.0004634
 \end{aligned}$$

Gambar 11. Perhitungan Nilai Tidak Puas

Kemudian kita akan menghitung nilai probabilitasnya dengan melakukan normalisasi terhadap *likelihood* tersebut sehingga jumlah nilai yang diperoleh menjadi satu, seperti pada gambar 12.

$$\text{Probabilitas Puas} = \frac{0.0014726}{0.0014726+0.0004634} = 0.76$$

$$\text{Probabilitas Tidak Puas} = \frac{0.0004634}{0.0014726+0.0004634} = 0.24$$

Gambar 12. Perhitungan Probabilitas Kepuasan Akhir

Sehingga bisa dilihat pada probabilitas diatas dapat diklasifikasikan bahwa pelanggan tersebut merasa PUAS saat berbelanja di toko tersebut dengan nilai probabilitas 76%.

### 2) Hasil Prediksi Menghitung Tingkat Kepuasan Menggunakan Program

Pada pengujian kali ini dengan menggunakan program AI dengan data yang sama seperti perhitungan manual sebelumnya. Sehingga didapatkan hasil prediksi menggunakan program sebagai berikut :

**Prediksi Naive Bayes : ['Puas']**

Gambar 12. Hasil Prediksi Tingkat Kepuasan Menggunakan AI

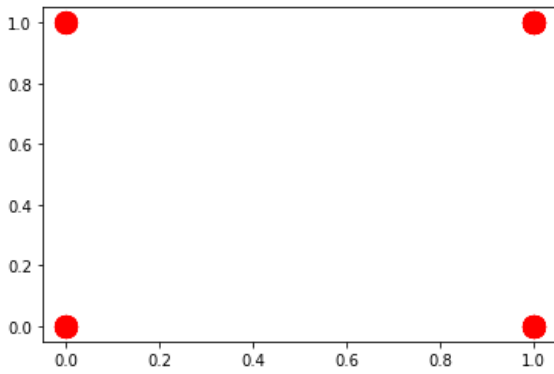
Hasil test dan prediksi dari program AI menunjukkan Tingkat Kepuasan pelanggan terhadap toko dengan atribut yang diinputkan menampilkan hasil prediksi PUAS. Sehingga dengan kondisi penilaian tersebut bisa dikategorikan bahwa pelanggan tersebut puas terhadap pelayanan toko tersebut.

### III. HASIL DAN PEMBAHASAN

#### 3.1. Memetakan Grafik Persebaran Data

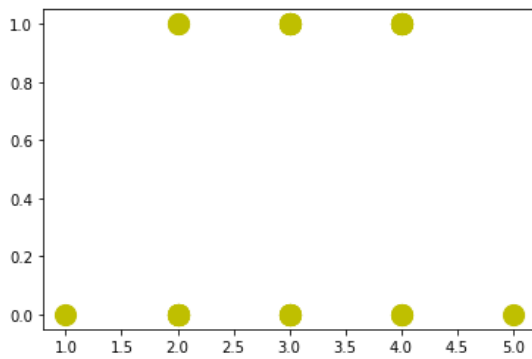
Pada *dataset* yang telah ada, dapat kita tampilkan grafik persebaran data untuk melihat keragaman data pada setiap komponen yang ada. Berdasarkan *dataset* yang ada dapat kita lihat persebaran data dalam grafik seperti dibawah ini :

##### 1) Persebaran data Komponen Harga



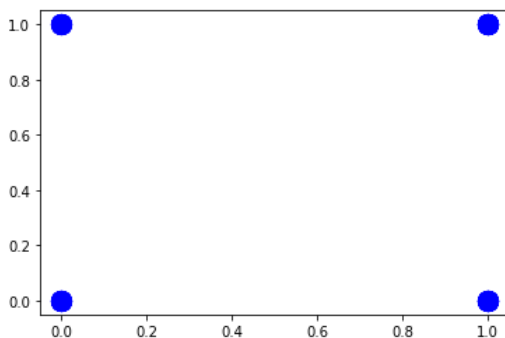
Gambar 13. Grafik persebaran data Harga

2) Persebaran data Komponen Keamanan



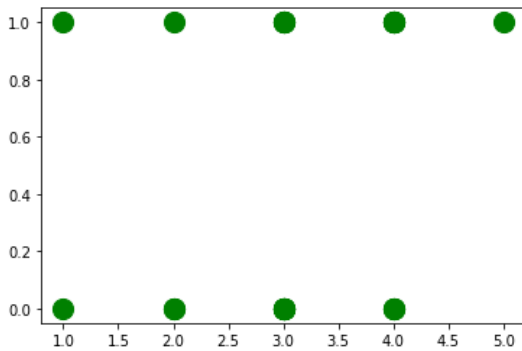
Gambar 14. Grafik persebaran data Keamanan

3) Persebaran data Komponen Lokasi



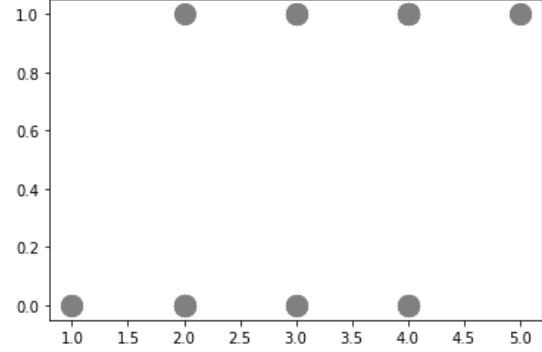
Gambar 15. Grafik persebaran data Lokasi

4) Persebaran data Komponen Fasilitas



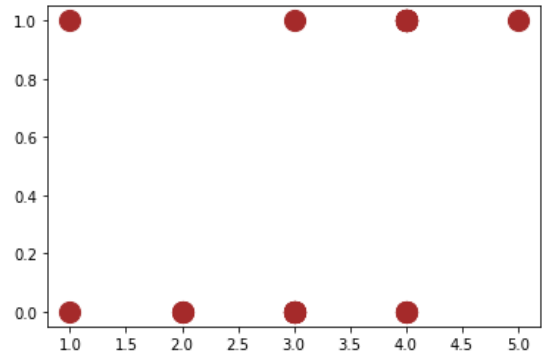
Gambar 16. Grafik persebaran data Fasilitas

5) Persebaran data Komponen Service



Gambar 17. Grafik persebaran data Service

6) Persebaran data Komponen Kelengkapan\_Barang



Gambar 18. Grafik persebaran data Kelengkapan\_Barang

3.2. Confusion Matrix Naïve Bayes

Pada dasarnya *confusion matrix* merupakan sebuah metode yang digunakan untuk menganalisa seberapa baik sebuah model dengan hasil klasifikasi sebenarnya. *Confusion matrix* berbentuk tabel matriks yang menggambarkan kinerja model klasifikasi pada serangkaian data uji yang nilai sebenarnya diketahui. Seperti akurasi, presisi dan juga *recall*, dengan perhitungan:

Tabel 1. Confusion Matrix

Hasil Aktual	Hasil Prediksi	
	Benar	Salah
Benar	TP	FN
Salah	FP	TN

Keterangan :

Hasil Aktual (Benar) : yaitu hasil aktual atau hasil sebenarnya yang bernilai benar.

Hasil Aktual (Salah) : yaitu hasil aktual atau hasil sebenarnya yang bernilai salah.

Hasil Prediksi (Benar) : merupakan hasil prediksi yang bernilai benar.

Hasil Prediksi (Salah): merupakan hasil prediksi yang bernilai salah.

TP (*True Positive*) : adalah banyaknya jumlah data yang di kategorikan (benar) pada hasil aktual juga bernilai (benar) pada hasil prediksi.

FN (*False Negative*): adalah banyaknya jumlah data yang di kategorikan (benar) pada hasil aktual namun bernilai (salah) pada hasil prediksi.

FP (*False Positive*): adalah banyaknya jumlah data yang di kategorikan (salah) pada hasil aktual namun bernilai (benar) pada hasil prediksi.

TN (*True Negative*): adalah banyaknya jumlah data yang di kategorikan (salah) pada hasil aktual juga bernilai (salah) pada hasil prediksi.

Setelah melakukan prediksi data terhadap *dataset* menggunakan Program *AI* didapatkan hasil data perbandingan antara data aktual dengan data prediksi sebagai berikut

No	Status Sebenarnya	Hasil Prediksi
1	Puas	Puas
2	Tidak Puas	Tidak Puas
3	Puas	Puas
4	Tidak Puas	Tidak Puas
5	Tidak Puas	Tidak Puas
6	Puas	Puas
7	Puas	Puas
8	Puas	Puas
9	Tidak Puas	Tidak Puas
10	Tidak Puas	Tidak Puas
11	Puas	Puas
12	Tidak Puas	Tidak Puas
13	Puas	Puas
14	Tidak Puas	Tidak Puas
15	Tidak Puas	Tidak Puas
16	Puas	Puas
17	Puas	Puas
18	Puas	Puas
19	Tidak Puas	Tidak Puas
20	Tidak Puas	Tidak Puas
21	Tidak Puas	Puas
22	Tidak Puas	Tidak Puas
23	Tidak Puas	Puas
24	Tidak Puas	Puas
25	Tidak Puas	Puas
26	Puas	Puas
27	Puas	Tidak Puas
28	Tidak Puas	Tidak Puas
29	Puas	Tidak Puas
30	Puas	Puas

Gambar 19. Tabel data aktual dan prediksi

Berdasarkan data dari tabel diatas diperoleh hasil *Confusion Matrix* sebagai berikut :

Tabel 2. Hasil *Confusion Matrix*

TP = 12	FP = 4
FN = 2	TN = 12

Setelah mengetahui hasil dari *Confusion Matrix* maka dapat kita hitung *Performance* dari perhitungan algoritma *Naïve Bayes* yang terdiri dari :

1. *Accuracy*

Akurasi menunjukkan tingkat rasio kebenaran suatu proses dalam menentukan dan mengelompokkan data

dengan benar.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Pada hasil penelitian di atas yang dihitung secara manual didapatkan hasil akurasi sebagai berikut :

$$accuracy = \frac{12 + 12}{12 + 2 + 4 + 12} = 80\%$$

Sehingga hasil akurasi yang didapatkan pada kasus diatas sebesar 80%.

2. *Precision*

*Precision* adalah tingkat rasio data aktual yang bernilai benar dengan hasil prediksi yang bernilai benar.

$$Precision = \frac{TP}{TP + FP}$$

Pada hasil penelitian di atas yang dihitung secara manual didapatkan hasil presisi sebagai berikut :

$$Precision = \frac{12}{12 + 4} = 75\%$$

Sehingga hasil Presisi yang didapatkan pada kasus diatas sebesar 75%.

3. *Recall*

*Recall* merupakan sebuah tingkat kesuksesan pada suatu sistem dalam mendapatkan kembali data aktual benar.

$$Recall = \frac{TP}{TP + FN}$$

Pada hasil penelitian di atas yang dihitung secara manual didapatkan hasil *recall* sebagai berikut:

$$Recall = \frac{12}{12 + 2} = 85,7\%$$

Sehingga hasil *recall* yang didapatkan pada kasus diatas sebesar 85,7%.

4. *F1 Score*

*F1 Score* adalah rasio rata-rata dari nilai presisi dan nilai *recall* yang didapatkan.

$$F1\ Score = 2 * (recall * precision) / (recall + precision)$$

Pada hasil penelitian di atas yang dihitung secara manual didapatkan hasil *f1 score* sebagai berikut :

$$F1\ Score = 2 * (85,7\% * 75\%) / (85,7\% + 75\%)$$

$$F1\ Score = 2 * (64,27\%) / (149,9\%)$$

$$F1\ Score = 85,75\%$$

Sehingga hasil *f1 score* yang didapatkan pada kasus diatas sebesar 85,75%.

Sedangkan pada perhitungan *Confusion Matrix* menggunakan program *AI* didapatkan hasil sebagai berikut :

	precision	recall	f1-score	support
0	0.86	0.75	0.80	16
1	0.75	0.86	0.80	14
accuracy			0.80	30
macro avg	0.80	0.80	0.80	30
weighted avg	0.81	0.80	0.80	30

Gambar 20. Hasil *Confusion Matrix* dengan Program *AI*

Dapat dilihat perhitungan *Confusion Matrix* menggunakan program *AI* relatif sama dengan hasil perhitungan manual.

**IV. KESIMPULAN**

*Naïve Bayes* merupakan metode paling sederhana dari pengklasifikasian probabilitas, yang memiliki tingkat akurasi sangat tinggi ketika diaplikasikan pada *database*

dengan *big data*. Terlebih lagi metode *Naïve Bayes* hanya memerlukan sedikit data pelatihan awal untuk mengestimasi parameter yang dibutuhkan untuk klasifikasi dan *Naïve Bayes* dikenal cepat dan efisiensi ruang penyimpanan.

Dari hasil penelitian yang dilakukan ini dapat kita peroleh hasil dari penerapan algoritma *Naïve Bayes* pada studi kasus yang dilakukan memiliki tingkat akurasi yang relatif bagus pada data yang besar namun tingkat *error* pada data harus kecil dan semakin beragam data maka akan semakin tinggi pula tingkat akurasinya.

Untuk penyempurnaan penelitian ini perlu dilakukan:

- 1) Jika menggunakan metode *Naïve Bayes*, *dataset* yang digunakan harus memiliki tingkat *error* yang rendah, agar tingkat akurasi pada data yang diprediksi memiliki rasio benar yang tinggi.
- 2) Metode *Naïve Bayes* memang memiliki tingkat akurasi yang tinggi, namun akan lebih baik apabila metode ini dikombinasikan dengan metode lain yang tepat agar bisa saling memperbaiki kelemahan yang ada.
- 3) Penggunaan metode *Naïve Bayes* akan lebih akurat apabila jumlah *dataset* yang digunakan berjumlah banyak, agar pada saat proses pelatihan data, tingkat akurasi pada data yang dicari jauh lebih akurat, karena memiliki banyak sampel untuk dipelajari.

#### IV. REFERENSI

- [1] M. V. Permana, "PENINGKATAN KEPUASAN PELANGGAN MELALUI KUALITAS PRODUK DAN KUALITAS LAYANAN," *J. Din. Manaj.*, vol. 4, no. 2, hal. 115–131, 2013.
- [2] E. Manalu, F. A. Sianturi, dan M. R. Manalu, "Penerapan Algoritma Naive Bayes Untuk Memprediksi Jumlah Produksi Barang Berdasarkan Data Persediaan dan Jumlah Pemesanan Pada CV. Papadan Mama Pastries," *J. Mantik Penusa*, vol. 1, no. 2, hal. 16–21, 2017, [Daring]. Tersedia pada: <https://ezp.lib.unimelb.edu.au/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=ffh&AN=2008-10-Aa4022&site=eds-live&scope=site>.
- [3] M. Siddik, Y. Desnelita, dan Gustientiedina, "Penerapan Naïve Bayes untuk Memprediksi Tingkat Kepuasan Mahasiswa Terhadap Pelayanan Akademis," *J. Infomedia*, vol. 2, no. 4, hal. 89–93, 2019, [Daring]. Tersedia pada: <http://e-jurnal.pnl.ac.id/index.php/infomedia/article/view/1892>.
- [4] Ali Hasan, "Jurnal Indovisi; Pengaruh Citra Merek, Kualitas Produk dan harga Terhadap Minat Beli Pakaian," *Indonesoan Indovisi Inst.*, vol. 1, no. Iii, hal. 83–104, 2018, [Daring]. Tersedia pada: <http://journal.dosenindonesia.org>.
- [5] H. Annur, "Klasifikasi Masyarakat Miskin Menggunakan Metode Naive Bayes," *Ilk. J. Ilm.*, vol. 10, no. 2, hal. 160–165, 2018, doi: 10.33096/ilkom.v10i2.303.160-165.