

## PENERAPAN METODE NAÏVE BAYES UNTUK PREDIKSI KEPUASAN PELANGGAN STUDI KASUS BENGKEL WIN MOTOR

Bambang Hermanto<sup>1)</sup>, Ahmad Romadhoni<sup>2)</sup>

Program Studi Teknik Informatika Sekolah Tinggi Teknologi Pelita Bangsa  
[bambang.hermanto@pelitabangsa.ac.id](mailto:bambang.hermanto@pelitabangsa.ac.id)

Disetujui, 27 Juni 2019

### Abstraksi

Ketatnya persaingan dibidang usaha perbengkelan membuat perusahaan yang bergerak dibidang jasa pelayanan atau perawatan kendaraan bermotor bersaing untuk menarik pelanggannya dalam membeli spare part dan jasa yang ditawarkan. Untuk mempertahankan para pelanggannya sebuah perusahaan harus dapat memahami secara seksama harapan – harapan apa saja dari pelanggannya sehingga perusahaan harus mengetahui tingkat kepuasan tiap pelanggannya. Pada Bengkel WIN MOTOR yang terletak diperumahan graha asri cikarang timur, kesulitan untuk menentukan apakah pelanggan merasa puas atau tidak dalam hal pelayanan yang diberikan. Dengan menggunakan teknik kuesioner, didapatkan hasil data kuesioner. Pengolahan data dengan menggunakan penerapan metode naïve bayes untuk memprediksi kepuasan pelanggan. Berdasarkan data hasil kuesioner yang diolah menggunakan metode naïve bayes untuk memprediksi kepuasan pelanggan di Bengkel WIN MOTOR didapatkan tingkat akurasi sebesar 90.00%. Setelah dilakukan pengujian kasus untuk menentukan hasil prediksi kepuasan pelanggan dengan menggunakan perhitungan manual dan menggunakan aplikasi Rapidminer, didapatkan hasil prediksi kepuasan pelanggan yang sama.

**Kata Kunci :** *Prediksi, Naïve Bayes, Kepuasan Pelanggan*

### Abstract

*The tight competition in the workshop business makes a company engaged in service or maintenance of motorized vehicles compete to attract its customers in buying spare parts and services offered. To maintain its customers, a company must be able to understand carefully the expectations - any expectations of its customers so that the company must know the level of satisfaction of each customer. At WIN MOTOR Workshop which is located in the residential area of East Cikarang Graha, it is difficult to determine whether the customer feels satisfied or not in terms of the services provided. By using questionnaire techniques, the results of questionnaire data are obtained. Processing data using the application of naïve Bayes methods to predict customer satisfaction. Based on the data from the questionnaire that was processed using the naïve bayes method to predict customer satisfaction in the WIN MOTOR Workshop, the accuracy rate was 90.00%. After testing cases to determine customer satisfaction predictions using manual calculations and using the Rapidminer application, the same customer satisfaction predication results were obtained.*

**Keywords :** *Prediction, Naïve Bayes, Customer Satisfaction*

### 1. Pendahuluan

Persaingan didunia usaha pada era globalisasi ini semakin ketat dari waktu ke waktu, baik pada skala kecil maupun skala yang cukup besar. Dalam hal ini pelaku usaha dituntut untuk bersaing dalam berbagai aspek. Salah satunya antara lain dalam hal kualitas pelayanan, kelengkapan fasilitas, harga dan lokasi yang paling tepat atau strategis. Hari ke hari semakin lama jumlah kendaraan bermotor terus bertambah, hal ini di sebabkan karena meningkatnya kebutuhan masyarakat akan kendaraan sebagai sarana transportasi yang sangat penting. Kendaraan tersebut pastilah membutuhkan suatu perawatan berkala demi memperoleh performa yang maksimal.

Sekarang ini bukan hanya jumlah kendaraan yang bertambah banyak, bengkel-bengkel motor pun sudah bertambah banyak tanpa disadari dan menambah ketatnya persaingan dibidang usaha pada perawatan kendaraan bermotor. Para konsumen pun bertambah jeli dalam memilih bengkel motor untuk memperbaiki kendaraan mereka. Disamping itu para pengusaha bengkel motor tidak hanya terpaku mengutamakan pelayanan yang baik, juga harus memperhatikan dari segi kelengkapan spare part, dan fasilitas bengkel yang memadai untuk tercapainya pelayanan yang memuaskan para konsumen. Menurut Philip Kotler dan Lane Keller yang dikutip dari buku Manajemen Pemasaran mengatakan bahwa kepuasan konsumen adalah perasaan senang atau kecewa

seseorang yang muncul setelah membandingkan kinerja (hasil) produk yang dipikirkan terhadap kinerja yang diharapkan yang dikutip oleh (Apriyani, Fakultas, & Administrasi, 2017).

Penulis bermaksud melakukan teknik mining untuk mengatasi masalah tersebut yaitu berupa sistem evaluasi kepuasan pelanggan yang nantinya akan berguna untuk membantu pengusaha bengkel Win Motor dalam menganalisa kepuasan pelanggan.

Metode Naïve Bayes yang akan nantinya digunakan oleh penulis untuk memprediksi kepuasan pelanggan. Naïve Bayes merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes. Metode Naïve Bayes memiliki tingkat akurasi yang sangat tinggi ketika diaplikasikan pada database dengan bigdata. Pada penelitian ini bertujuan untuk mengklasifikasikan kepuasan terhadap pelayanan yang diberikan bengkel WIN Motor dengan menggunakan metode Naïve Bayes, untuk mengetahui seberapa besar tingkat akurasi dalam membuat klasifikasi kepuasan pelanggan. Agar nantinya dapat membantu bengkel Win Motor mengevaluasi dan dapat meningkatkan kualitas pelayanan supaya menjadi bengkel motor yang selalu menjadi pilihan para konsumen

## **2. Tinjauan Studi**

### **2.1. Definisi Prediksi**

Prediksi adalah suatu proses memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi di masa depan berdasarkan informasi masa lalu dan sekarang yang dimiliki, agar kesalahannya (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat diperkecil. Prediksi tidak harus memberikan jawaban secara pasti kejadian yang akan terjadi, melainkan berusaha untuk mencari jawaban sedekat mungkin / setepat mungkin yang akan terjadi. Dikutip oleh (Apriyani et al., 2017).

Pengertian Prediksi sama dengan ramalan atau perkiraan. Menurut kamus besar bahasa Indonesia, prediksi adalah hasil dari kegiatan memprediksi atau meramal atau memperkirakan nilai pada masa yang akan datang dengan menggunakan data masa lalu. Prediksi menunjukkan apa yang akan terjadi pada suatu keadaan tertentu dan merupakan input bagi proses perencanaan dan pengambilan keputusan.

### **2.2. Definisi Kepuasan Pelanggan**

Kepuasan menurut (Laksana 2008, 9) adalah harapan sama dengan kenyataan. Kemudian menurut Kotler dalam (Laksana 2008, 9) yang dikutip oleh (Fauzi et al, 2016) kepuasan pelanggan yaitu: "Customer satisfaction is the outcome felt by buyer who have experienced a company performance that has fulfilled expectations". Maksudnya yaitu menyangkut komponen harapan dan kinerja/hasil yang dirasakan. Pada umumnya harapan pelanggan merupakan perkiraan atau keyakinan pelanggan tentang apa yang akan diterima apabila ia membeli atau mengkonsumsi suatu produk baik barang maupun jasa, sedangkan kinerja atau hasil yang dirasakan merupakan persepsi pelanggan terhadap apa yang ia terima setelah mengkonsumsi produk yang ia beli. Untuk itu manajemen harus memiliki persepsi yang sama dengan pelanggan agar supaya diperoleh hasil yang melebihi atau paling tidak sama dengan harapan pelanggan. Keadaan ini sama seperti yang telah dikemukakan menurut Gasperz dalam (Laksana 2008,10) yang dikutip oleh (Warmadewa & Suryani, 2014) "Kepuasan pelanggan dapat didefinisikan secara sederhana sebagai suatu keadaan dimana kebutuhan, keinginan dan harapan pelanggan dapat terpenuhi melalui produk yang dikonsumsi".

### **2.3. Data Mining**

Data mining merupakan suatu proses untuk mengidentifikasi pola yang memiliki potensi dan berguna untuk mengelola dataset yang besar (Witten & Frank, 2011). Dalam data mining ada 10 algoritma teratas yang paling berpengaruh yang dipilih oleh peneliti dalam komunitas data mining, dimana 6 (enam) diantaranya adalah algoritma klasifikasi yaitu C4.5, Support Vector Machines (SVM), AdaBoost, k Nearest Neighbor (kNN), Naïve Bayes dan CART (Fayyad, Piatetsky-Shapiro, & Smyth, 1996).

Data mining adalah suatu proses yang menggunakan statistical, mathematical, artificial intelligence dan machine learning techniques untuk mengekstrak dan mengidentifikasi informasi penting dan subsequence knowledge dari database yang besar. Dimana informasi yang diekstrak digunakan dalam mengambil keputusan bisnis yang cukup krusial (crucial business decision).

### **2.4. Fungsi Data Mining**

Data Mining memiliki kemampuan cepat dalam melakukan analisis dan sangat focus pada variabel-variabel penting. Patterns (pola) dan rules dalam (aturan, kaidah) yang dapat digunakan dalam membuat keputusan dan forecast (meramalkan) dampak dari keputusan tersebut. Intelligent data mining meliputi informasi dalam data warehouse dimana query dalam laporan biasa tidak bisa mengungkapkan informasi secara efektif. Data mining tools mampu menemukan pola dalam data dan memberi rules (Jananto, 2013).

Ada tiga metode yang digunakan dalam mengidentifikasi pola dalam data yaitu: simple model (contoh, SQL berbasis query yang merupakan cara sederhana dalam menarik data, online analytical processing (OLAP), human judgement); Intermediate models (contoh, regression, decision, trees clustering); complex models (contoh, neural networks, other rule induction). Algoritma data mining tradisional juga membagi empat kategori besar yakni classification, regression, anomaly detection, time series, clustering dan association analysis.

**2.5. Model Data Mining**

Ada berbagai model dalam data mining atau sering disebut teknik data mining, secara umum model data mining dibagi dalam tiga kelompok berdasarkan pada tugas atau fungsi yang terdiri dari classification, clustering dan associaton.

Klasifikasi melakukan analisa pada data sejarah yang tersimpan dalam database dan meregenerate otomatis model yang dapat memprediksikan masa depan. Dengan melakukan redefined class, model dapat memprediksi sebuah kelas atau membuat kelas pada record-record data yang terklasifikasi. Klasifikasi menemukan pola data yang digunakan untuk mengklasifikasi dalam kategori tertentu, contohnya pada aplikasi email yang dapat mengklasifikasi email yang bukan spam dan email spam. Contoh lain klasifikasi antara pelanggan yang membeli produk terbanyak dengan pelanggan membeli produk dalam jumlah sedikit. Informasi ini misalnya bisa digunakan dalam melakukan iklan, tentu iklan akan di fokuskan kepada pelanggan yang memiliki jumlah pembelian paling banyak, karena besar peluang pelanggan tersebut untuk membeli kembali. Algoritma yang biasa digunakan dalam klasifikasi adalah neural network, decision trees, naïve bayes, dan if then else rules (Jiawei, H., Kamber, M., Han, J., Kamber, M., & Pei, J. 2012. P.327) yang dikutip oleh (Jadav & Panchal, 2012).

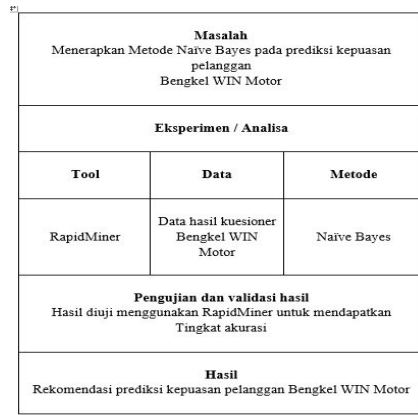
**2.6. Naïve Bayes**

Algoritma Naïve Bayes akan mengevaluasi setiap atribut yang berkontribusi prediksi pada atribut target. Naïve Bayes tidak memperhitungkan relasi antar atribut-atribut contributor prediksi, tidak seperti Decision Tree yang memperhitungkan relasi antara atribut. Bentuk tugas dasar yang dilakukan oleh algoritma Naïve Bayes adalah hanyalah klasifikasi (Jiawei, H., Kamber, M., Han, J., Kamber, M., & Pei, J.2012.p.351) yang dikutip oleh (Degree, Science, & Lecture, 2012). Naïve bayes merupakan teknik data mining dengan pendekatan teori probabilitas untuk membangun sebuah model klasifikasi berdasarkan pada kejadian masa lalu yang mempunyai potensi membentuk sebuah objek baru yang dikategorikan sebagai kelas yang memiliki probabilitas terbaik.

Naïve Bayes memiliki kemampuan yang cepat dalam membuat model, mempunyai kemampuan yang cepat dalam menyediakan metode baru dalam mengeksplor dan memahami data

**3. Kerangka Konsep**

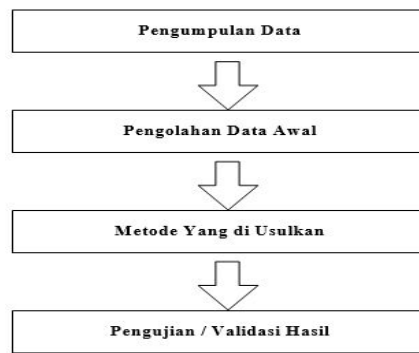
Kerangka konsep dalam penelitian ini secara sistematis dan sederhana dapat gambarkan seperti berikut ini:



**Gambar 1. Kerangka Konsep**

**4. Desain Penelitian/Methodologi**

Pada penelitian ini, data yang digunakan adalah data hasil kuesioner pada bengkel Win Motor data tersebut akan diolah menggunakan beberapa metode data mining sehingga diperoleh satu metode yang paling akurat dan dapat digunakan sebagai rules dalam memprediksi kepuasan pelanggan. Berikut tahapan dalam melakukan penelitian data mining:



**Gambar 2.** Tahapan Penelitian

Pada tahapan ini menjelaskan tentang bagaimana dan dari mana sumber data didapatkan. Data sekunder berisi tentang sumber perolehan data untuk keperluan penelitian, sedangkan data primer berisi tentang data yang dihasilkan dari penelitian.

Pada tahapan pengolahan data awal menjelaskan tentang awal data mining. Data yang telah didapatkan akan diolah keformat yang dibutuhkan pengelompokkan dan penentuan atribut data.

Pada tahap metode yang diusulkan untuk digunakan pada prediksi kepuasan pelanggan. Penjelasan meliputi pengaturan pemilihan nilai dari parameter parameter dan arsitektur melalui uji coba.

Pada tahap pengujian metode menjelaskan tentang langkah-langkah eksperimen meliputi pemilihan arsitektur yang tepat dari model atau metode yang diusulkan sehingga didapatkan hasil yang dapat membuktikan bahwa metode yang digunakan adalah tepat.

#### 4.1. Metode Pengumpulan Data

Metode pengumpulan data yang dilakukan penulis untuk mendapatkan data yang valid adalah sebagai berikut :

1. Metode kuesioner, metode kuesioner dilakukan dengan cara memberikan form kuesioner terhadap pelanggan bengkel WIN Motor untuk mendapatkan data kepuasan pelanggan.
2. Metode Studi literatur, penulis mengumpulkan, membaca, mempelajari dan mencatat literatur dari jurnal maupun buku yang berkaitan dengan metode Naïve Bayes. Dari metode ini diharapkan dapat memberi

Data yang akan digunakan dalam penelitian ini adalah mengumpulkan data form quisioner yang telah dikumpulkan atau didapatkan dan diubah menjadi data excel. Jumlah data yang akan digunakan adalah 500 data pelanggan bengkel Win Motor.

**Tabel 1.** Data Pelanggan

Nama Pelanggan	Pelayanan	Respon	Garansi	Komunikasi	Fasilitas	Harga	Hasil
Saman Bin Narin	Sesuai	Cepat	Cukup	Mudah	Lengkap	Murah	Puas
Widiyah	Sesuai	Cepat	Cukup	Mudah	Kurang	Murah	Puas
Tiah Setiani	Sesuai	Cepat	Cukup	Sulit	Lengkap	Murah	Puas
Nur Elbadriah	Sesuai	Cepat	Cukup	Sulit	Kurang	Murah	Tidak
Yendri Rahmat	Sesuai	Cepat	Cukup	Mudah	Lengkap	Murah	Puas
Dede Susanti	Sesuai	Cepat	Kurang	Sulit	Lengkap	Murah	Tidak
Andhika Arifa Putra	Sesuai	Cepat	Kurang	Mudah	Kurang	Murah	Puas
Bangkit Prihatanto	Sesuai	Cepat	Kurang	Sulit	Kurang	Murah	Tidak
Pukim Herniawan	Sesuai	Lambat	Cukup	Mudah	Lengkap	Murah	Puas
Kartubi	Sesuai	Lambat	Cukup	Mudah	Kurang	Murah	Puas
Heni	Sesuai	Lambat	Cukup	Sulit	Lengkap	Murah	Puas
Budi Hermawan	Sesuai	Lambat	Cukup	Sulit	Kurang	Murah	Tidak
Sutisna	Sesuai	Lambat	Kurang	Mudah	Lengkap	Murah	Puas

Trisni Wahyuningsih	Sesuai	Lambat	Kurang	Sulit	Lengkap	Murah	Tidak
Eri Kristianis	Sesuai	Lambat	Kurang	Mudah	Kurang	Murah	Tidak
Ozi Syahputra	Sesuai	Lambat	Kurang	Sulit	Kurang	Murah	Tidak
Nur Agustian	Kurang	Cepat	Cukup	Mudah	Lengkap	Murah	Puas
Sopiandi	Kurang	Cepat	Cukup	Mudah	Kurang	Murah	Puas
Nurcahyati	Kurang	Cepat	Cukup	Sulit	Lengkap	Murah	Puas
Rizki Amalia Fajrin	Kurang	Cepat	Cukup	Sulit	Kurang	Murah	Tidak
Joni Purnawan	Kurang	Cepat	Kurang	Mudah	Lengkap	Murah	Puas
Tulistio	Kurang	Cepat	Kurang	Sulit	Lengkap	Murah	Tidak
Ayu Widhaningsih	Kurang	Cepat	Kurang	Mudah	Kurang	Murah	Tidak
Tohari	Kurang	Cepat	Kurang	Sulit	Kurang	Murah	Tidak
Jamal	Kurang	Lambat	Cukup	Mudah	Lengkap	Murah	Tidak
Agus Suprianto Sinaga	Kurang	Lambat	Cukup	Mudah	Kurang	Murah	Tidak

#### 4.2. Pengelolaan Data Awal

Dalam proses pengolahan data awal, beberapa tahapan akan dilakukan agar didapatkan data yang dapat diproses oleh penulis. Berikut proses tahapan tahapannya yaitu :

##### 1. Select Data

Data yang ada akan dilakukan pemilihan terhadap atribut / variabel yang akan dianalisis. Pemilihan atribut / variabel tersebut dengan mempertimbangkan tujuan penulis, sehingga akan didapatkan beberapa atribut / variable yang akan digunakan untuk menjadi masukan atau variabel input.

**Tabel 2.** Pemelihan Data

No	Atribut	Indikator	Datail Penggunaan
1	Nama Pelanggan	x	-
2	Pelayanan	√	Nilai Model
3	Respon	√	Nilai Model
4	Garansi	√	Nilai Model
5	Komunikasi	√	Nilai Model
6	Fasilitas	√	Nilai Model
7	Harga	√	Nilai Model
8	Hasil	√	Label Target

Tabel diatas menerangkan atribut yang akan dipakai dalam penelitian ini. Indikator “√” menandakan atribut akan digunakan, sedangkan indicator “x” menandakan atribut tersebut akan di eliminasi pada tahap pengolahan data awal.

##### 2. Cleaning

Pada tahap ini akan dilakukan proses pembersihan data untuk memastikan data yang telah dipilih tersebut telah layak untuk dilakukan proses pemodelan. Tahapan ini antara lain memperbaiki data yang rusak, membersihkan dan menghapus data yang tidak diperlukan. Pada tahap ini dilakukan proses menghilangkan atribut nama karena tidak digunakan pada proses pemodelan.

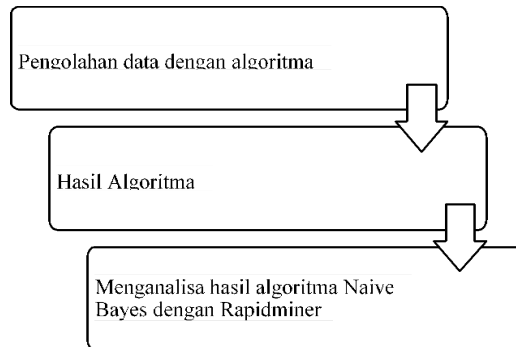
**Tabel 3.** Proses Cleaning Data

Pelayanan	Respon	Garansi	Komunikasi	Fasilitas	Harga	Hasil
Sesuai	Cepat	Cukup	Mudah	Lengkap	Murah	Puas
Sesuai	Cepat	Cukup	Mudah	Kurang	Murah	Puas
Sesuai	Cepat	Cukup	Sulit	Lengkap	Murah	Puas
Sesuai	Cepat	Cukup	Sulit	Kurang	Murah	Tidak
Sesuai	Cepat	Cukup	Mudah	Lengkap	Murah	Puas
Sesuai	Cepat	Kurang	Sulit	Lengkap	Murah	Tidak
Sesuai	Cepat	Kurang	Mudah	Kurang	Murah	Puas
Sesuai	Cepat	Kurang	Sulit	Kurang	Murah	Tidak
Sesuai	Lambat	Cukup	Mudah	Lengkap	Murah	Puas

Sesuai	Lambat	Cukup	Mudah	Kurang	Murah	Puas
Sesuai	Lambat	Cukup	Sulit	Lengkap	Murah	Puas
Sesuai	Lambat	Cukup	Sulit	Kurang	Murah	Tidak
Sesuai	Lambat	Kurang	Mudah	Lengkap	Murah	Puas
Sesuai	Lambat	Kurang	Sulit	Lengkap	Murah	Tidak
Sesuai	Lambat	Kurang	Mudah	Kurang	Murah	Tidak
Sesuai	Lambat	Kurang	Sulit	Kurang	Murah	Tidak
Kurang	Cepat	Cukup	Mudah	Lengkap	Murah	Puas
Kurang	Cepat	Cukup	Mudah	Kurang	Murah	Puas
Kurang	Cepat	Cukup	Sulit	Lengkap	Murah	Puas
Kurang	Cepat	Cukup	Sulit	Kurang	Murah	Tidak

3. Pemodelan

Dalam penelitian ini akan dilakukan analisa menggunakan metode klasifikasi dengan algoritma Naïve bayes. Data dihitung dengan metodenya kemudian di cari hasil akurasi. Dalam tahapan ini akan dilakukan beberapa langkah pengujian data yaitu seperti berikut.



**Gambar 3.** Langkah Pengujian Metode

Data Pelanggan akan diolah dengan algoritma Naïve Bayes, kemudian dilihat tingkat akurasi dari algoritma ini. Algoritma Naïve bayes akan mengevaluasi setiap atribut yang berkontribusi prediksi pada atribut target. Naïve Bayes tidak memperhitungkan relasi antar atribut-atribut kontributor prediksi, tidak seperti relasi antara atribut. Tugas dasar algoritma Naïve bayes adalah hanyalah klasifikasi.

4. Evaluasi

Melakukan pengecekan terhadap setiap nilai atribut dan model yang sudah dibangun, kemudian melakukan evaluasi terhadap hasil dengan melakukan analisis data mining. pada tahap ini juga merupakan tahapan dimana dilakukan perbaikan kembali bila terjadi kekurangan pada tahapan ini bisa saja kembali lagi ke tahap yang pertama dan kemudian ke tahap berikutnya dengan tujuan perbaikan, sampai sesuai dengan kebutuhan

5. Variable Penelitian

Variabel penelitian yang akan digunakan sebagai atribut data untuk proses data mining ialah pelayanan, respon, garansi, komunikasi, fasilitas dan harga. Kelas hasil adalah variabel target penelitian yang berisi 2 nilai kelas yaitu “ Puas “ dan “ Tidak “.

**4.3. Software Pengujian Data Mining**

Dalam melakukan proses data mining ini, perangkat Lunak yang akan digunakan Rapidminer Sebagai perangkat lunak untuk melakukan Analisa, proses data *mining*, *modelling*, *validasi*, akurasi model dan prediksi.

**5. Hasil Penelitian Dan Pembahasan**

**5.1. Hasil Penelitian**

1. Perhitungan *Naïve Bayes*

Pada tahap persiapan data bab ini adalah tahapan lanjutan persiapan data di bab sebelumnya. Berikut adalah tahapan lanjutannya.

a. Implementasi Algoritma Naïve bayes

Data yang digunakan adalah data training 100 record yang sudah ditentukan hasil keputusannya dan 20 data testing untuk mengimplementasikan algoritma naïve bayes.

b. Menghitung Probabilitas Kelas

Tahap pertama perhitungan untuk menentukan kepuasan pelanggan dengan metode naïve bayes adalah dengan mencari probabilitas dari masing masing kelas. Dalam penentuan kepuasan pelanggan akan ditentukan 2 kelas yaitu kelas "Puas" dan "Tidak". Cara perhitungannya adalah dengan mencari berapa jumlah data yang ada pada kelas “Puas” dan berapa jumlah pada kelas “Tidak” dari total keseluruhan data training, lalu membaginya dengan total keseluruhan data. Hasil dari perhitungan dapat dilihat pada tabel di bawah ini:

**Tabel 4.** Probabilitas Kelas

- P ( Puas )	=	31/100	Jumlah data "Puas" pada kolom "Hasil" di bagi jumlah data
- P( Tidak )	=	69/100	Jumlah data "Tidak" pada kolom "Hasil" di bagi jumlah data

- c. Mengitung Probabilitas Masing-Masing Atribut  
 Cara mencari probabilitas suatu atribut adalah dengan membandingkan atribut dari data testing dengan atribut dari data training. Berapa jumlah atribut dengan kelas "Puas" yang berada pada data training kemudian bagi dengan probabilitas kelas "Puas". Begitu juga dengan mencari probabilitas untuk kelas "Tidak".
- Pelayanan

**Tabel 5.** Atribut Pelayanan

Pelayanan	PUAS	Tidak	P ( PUAS )	P ( TIDAK )
Sesuai	27	8	27/31	8/69
Kurang	4	28	4/31	28/69
Tidak	0	33	0/31	33/69
Total	31	69	100%	100%

- Respon

**Tabel 6.** Atribut Respon

Respon	PUAS	Tidak	P ( PUAS )	P ( Tidak )
cepat	20	31	20/31	31/69
lambat	11	38	11/31	38/69
Total	31	69	100%	100%

- Garansi

**Tabel 7.** Atribut Garansi

Garansi	PUAS	Tidak	P ( PUAS )	P ( Tidak )
cukup	21	31	21/31	31/69
kurang	10	38	10/31	38/69
Total	31	69	100%	100%

- Komunikasi

**Tabel 8.** Atribut Komunikasi

Komunikasi	PUAS	Tidak	P ( PUAS )	P ( Tidak )
mudah	21	30	21/31	30/69
sulit	10	39	10/31	39/69
Total	31	69	100%	100%

- Fasilitas

**Tabel 9.** Atribut Fasilitas

Fasilitas	PUAS	Tidak	P ( PUAS )	P ( Tidak )
lengkap	20	32	20/31	32/69
kurang	11	37	11/31	37/69
Total	31	69	100%	100%

- Harga

**Tabel 10.** Atribut Harga

Harga	PUAS	Tidak	P ( PUAS )	P ( Tidak )
murah	15	35	15/31	35/69
mahal	16	34	16/31	34/69
Total	31	69	100%	100%

d. Mengitung Probabilitas Akhir Untuk Setiap Kelas  
 Untuk menghitung probabilitas akhir pada setiap kelas, perlu menggunakan data training dan mengubahnya menjadi nilai yang sudah ditentukan pada proses 3 sesuai atribut masing-masing. Lalu dari masing- masing atribut dan nilai probabilitas kelas dikalikan. Dari kedua hasil yang sudah ditentukan pada tiap kelas. bandingkan nilai yang paling tinggi. Jika kelas "Puas" bernilai paling tinggi maka hasilnya "Puas" begitu pula sebaliknya.

e. Perhitungan Naïve Bayes

Tahap I – Menghitung jumlah kelas/label keputusan :

- P( Di Angkat ) = 31/100 = 0,31
Jumlah data "Puas" pada kolom "Hasil" di bagi jumlah data
- P( Tidak ) = 69/100 = 0,69
Jumlah data "Tidak" pada kolom "Hasil" di bagi jumlah data

Tahap II – Menghitung jumlah kasus yang sama kelas dengan kelas yang sama :

P (   Hasil = Puas )		31
P ( Pelayanan	Sesuai	27 = 27/31 = 0.8710
P ( Respon	Lambat	11 = 11/31 = 0.3548
P ( Garansi	Kurang	10 = 10/31 = 0.3226
P ( Komunikasi	Mudah	21 = 21/31 = 0.6774
P ( Fasilitas	Lengkap	20 = 20/31 = 0.6452
P ( Harga	Murah	15 = 15/31 = 0.4839
P( Hasil = Puas )		31/100 = 0.3100
Mengalikan semua hasil atribut Puas =		6.5356

P (   Hasil = Tidak )		69
P ( Pelayanan	Sesuai	8 = 8/69 = 0.1159
P ( Respon	Lambat	38 = 38/69 = 0.5507
P ( Garansi	Kurang	38 = 38/69 = 0.5507
P ( Komunikasi	Mudah	30 = 30/69 = 0.4348
P ( Fasilitas	Lengkap	32 = 32/69 = 0.4638
P ( Harga	Murah	15 = 35/69 = 0.2174
P( Hasil = Tidak )		69/100 = 0.6900



Mengalikan semua hasil atribut Tidak	= 1.0636
--------------------------------------	----------

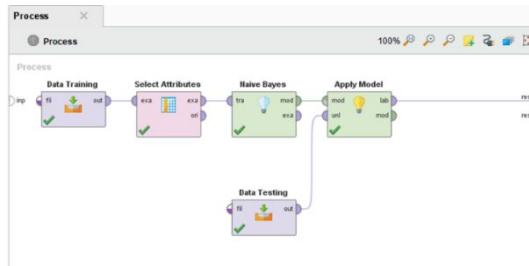
Tahap III – Membandingkan nilai kelas puas dan tidak puas

$P(\text{Keputusan} = \text{Puas}   x)$	= 6.5356
$P(\text{Keputusan} = \text{Tidak}   x)$	= 1.0636

Hasil $(P PUAS) = 6.5356$ lebih besar dari $(P TIDAK) = 1.0636$ maka hasilnya adalah "PUAS".
--

2. Analisa *Algoritma Naïve Bayes* dengan *Rapidminer*

Rapidminer 7.6 adalah salah satu tool untuk melakukan prediksi dan analisa data mining. Berikut adalah uji coba yang dilakukan pada 20 data testing yang sudah ditentukan. Data testing tersebut akan dicari nilai prediksinya apakah sesuai dengan perhitungan naïve bayes.



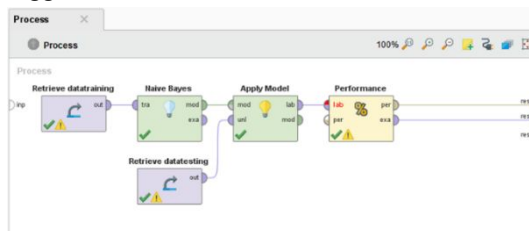
Gambar 4. Proses Analisa Dengan *Rapidminer*

predic...	confidence...	confidence...	PELAYANAN	RESPON	GARANSI	KOMUNIKASI	FASILITAS	HARGA
TIDAK	0.000	1.000	TIDAK	CEPAT	KURANG	SULIT	KURANG	MURAH
TIDAK	0.000	1.000	TIDAK	LAMBAT	KURANG	MUDAH	LENGKAP	MURAH
TIDAK	0.250	0.750	KURANG	CEPAT	CUKUP	MUDAH	KURANG	MAHAL
TIDAK	0.115	0.885	KURANG	CEPAT	KURANG	MUDAH	KURANG	MAHAL
TIDAK	0.052	0.948	KURANG	LAMBAT	CUKUP	SULIT	KURANG	MAHAL
TIDAK	0.001	0.999	TIDAK	CEPAT	CUKUP	MUDAH	LENGKAP	MAHAL
TIDAK	0.000	1.000	TIDAK	CEPAT	KURANG	SULIT	LENGKAP	MAHAL
TIDAK	0.000	1.000	TIDAK	LAMBAT	CUKUP	SULIT	LENGKAP	MAHAL
TIDAK	0.000	1.000	TIDAK	LAMBAT	KURANG	SULIT	KURANG	MAHAL
PUAS	0.938	0.062	SERLIAN	CEPAT	CUKUP	MUDAH	LENGKAP	MURAH
PUAS	0.682	0.318	SERLIAN	CEPAT	KURANG	SULIT	LENGKAP	MURAH
PUAS	0.763	0.237	SERLIAN	LAMBAT	CUKUP	MUDAH	KURANG	MURAH
PUAS	0.556	0.444	SERLIAN	LAMBAT	KURANG	MUDAH	KURANG	MURAH
PUAS	0.887	0.113	SERLIAN	CEPAT	CUKUP	MUDAH	KURANG	MAHAL
PUAS	0.529	0.471	SERLIAN	CEPAT	KURANG	SULIT	KURANG	MAHAL
PUAS	0.724	0.276	SERLIAN	LAMBAT	KURANG	MUDAH	LENGKAP	MURAH

Gambar 5. Hasil Pengujian Menggunakan *Rapidminer*

3. Akurasi Prediksi

Proses klasifikasi dengan Rapidminer menggunakan metode naïve bayes yang digunakan untuk mengklasifikasi data pelanggan.



Gambar 6. Proses *Training & Testing* Menggunakan *Rapidminer*

5.2. Hasil Pengujian

Dengan mengetahui jumlah data yang diklasifikasikan secara benar maka dapat diketahui akurasi hasil prediksinya 90.00% dari hasil data testing :

accuracy: 90.00%

	true PUAS	true TIDAK	class precision
pred. PUAS	5	2	71.43%
pred. TIDAK	0	13	100.00%
class recall	100.00%	86.67%	

Gambar 7. Akurasi Hasil Testing

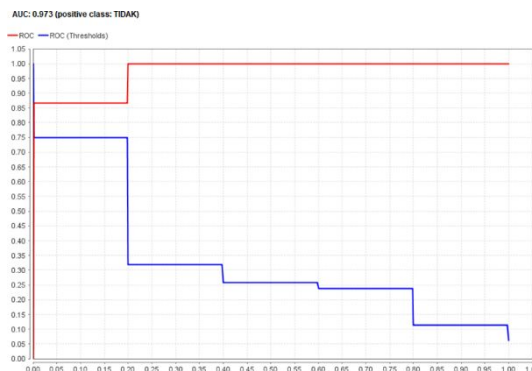
Untuk menghitung akurasinya seperti berikut :  
 Jumlah data yang diuji 20

Jumlah data yang diprediksi benar "Puas"	5
Jumlah data yang diprediksi benar "Tidak"	13
Jumlah data yang diprediksi salah "Puas"	0
Jumlah data yang diprediksi salah "Tidak"	2

$$\begin{aligned} \text{Akurasi} &= \text{Jumlah data yang diprediksi benar} / \text{jumlah data yang diuji} * 100\% \\ &= (5+13) / (5+13+0+2) * 100\% \\ &= 90\% \end{aligned}$$

$$\begin{aligned} \text{Error} &= \text{Jumlah data yang diprediksi salah} / \text{jumlah data yang diuji} * 100\% \\ &= (0+2) / (5+13+0+2) * 100\% \\ &= 10\% \end{aligned}$$

Dari perhitungan tersebut dapat disimpulkan bahwa klasifikasi dengan menggunakan metode Naive Bayes untuk menentukan kepuasan pelanggan menghasilkan tingkat akurasi sebesar 90% dan tingkat error 10%.



**Gambar 8.** Kurva ROC

Kurva ROC digunakan untuk menekspressikan data. Garis horizontal mewakili nilai false positive dan garis vertical mewakili nilai true positive. Dari gambar 4.5 dapat diketahui bahwa nilai Area Under Curve (AUC) model algoritma naïve bayes 0.973, hal ini menunjukkan bahwa model algoritma naïve bayes mencapai klasifikasi hampir sempurna.

## 6. Kesimpulan

Setelah melakukan penelitian penerapan metode naïve bayes untuk prediksi kepuasan pelanggan pada studi kasus bengkel Win Motor dapat disimpulkan beberapa hal sebagai berikut:

1. Dari hasil penelitian yang telah dilakukan, didapatkan hasil prediksi menentukan kepuasan pelanggan dengan menggunakan tools Rapidminer dan juga perhitungan manual dengan menghitung probabilitas kelas. Setelah dilakukan pengujian terhadap data training dan data testing dengan menggunakan aplikasi Rapidminer didapatkan tingkat akurasi sebesar 90%.
2. Setelah dilakukan pengujian kasus untuk menentukan hasil prediksi kepuasan pelanggan dengan menggunakan perhitungan manual dan menggunakan aplikasi Rapidminer, didapatkan hasil prediksi kepuasan pelanggan yang sama.

## Daftar Pustaka

- Apriyani, D. A., Fakultas, S., & Administrasi, I. (2017). PENGARUH KUALITAS PELAYANAN TERHADAP KEPUASAN KONSUMEN (Survei pada Konsumen The Little A Coffee Shop Sidoarjo). *Jurnal Administrasi Bisnis*, 51(2), 1–7.
- Degree, M. M., Science, C., & Lecture, A. C. (2012). *Data Mining: Concepts and*, 5, 703. <https://doi.org/10.1093/nar/gku1019>
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazine*, 37–54. <https://doi.org/10.1145/240455.240463>
- Hamta, F., Tetap, D., Akuntansi, P., Ekonomi, F., Riau, U., Batam, K., ... Pendahuluan, I. (n.d.). *ANALISIS PENERAPAN DATA MINING DALAM MENGIKUR*, 1–17.
- Ilmiah, D. K., & Informatika, P. T. (2016). *Dokumen Karya Ilmiah | Skripsi | Prodi Teknik Informatika - S1 | FIK | UDINUS | 2016*, 0–1.
- Jadav, J., & Panchal, M. (2012). Association Rule Mining Method On OLAP Cube. *International Journal of Engineering Research and Technology*, 2(2), 1147– 1151. Retrieved from <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Association+Rule+Mining+Method+On+OLAP+Cube#0>

- Jananto, A. (2013). Algoritma Naive Bayes untuk Mencari Perkiraan Waktu Studi Mahasiswa. *Teknologi Informasi*, 18(1), 9–16.
- Malik, H. H., & Bhardwaj, V. S. (2011). Automatic training data cleaning for text classification. *Proceedings - IEEE International Conference on Data Mining, ICDM*, 442–449. <https://doi.org/10.1109/ICDMW.2011.36>
- McCranie, K. D., Faulkner, M., French, D., Daddis, G. A., Gow, J., & Long, A. (2011). <No Title>. *Journal of Strategic Studies*, 34(2), 281–293. <https://doi.org/10.1080/01402390.2011.569130>
- Warmadewa, P. U., & Suryani, N. K. (2014). PERAN CUSTOMER SATISFACTION SEBAGAI MEDIATOR, 12, 84–91.