

DATA MINING TECHNIQUE USING NAÏVE BAYES ALGORITHM TO PREDICT SHOPEE CONSUMER SATISFACTION AMONG MILLENNIAL GENERATION

Margaretha Intan Pratiwi Hant^{*1}, Hendry²

^{1,2}Program Studi Teknik Informatika, Universitas Kristen Satya Wacana, Indonesia
Email: ¹672018126@student.uksw.edu, ²hendry@uksw.edu

(Naskah masuk: 25 April 2022, Revisi : 06 Mei 2022, diterbitkan: 20 Agustus 2022)

Abstract

Shopee is one of the largest e-commerce platforms currently being used by Millennials. The use of Shopee itself makes it very easy for consumers to process transactions. Shopee itself is committed to maintaining and improving customer satisfaction so they don't switch to other competitors. However, it is undeniable that there are still many cases that can harm consumers when using the platform. With the cases that occur, it is very possible that there will be a big influence on the level of consumer satisfaction on the platform. Consumers will feel satisfied when the product or service used can meet consumer expectations. This study was made with the aim of predicting the level of consumer satisfaction of Shopee Indonesia among the Millennial Generation. This study applies data mining using the Naive Bayes Algorithm. The Naive Bayes algorithm itself is a simple probability classification that can calculate all possibilities by combining a number of combinations and the frequency of a value from the database obtained. The attributes used in conducting this research include Name, Gender, Age, Price, Performance and Efficiency, Fulfillment, Reliability, Control and Security, Responsiveness, Compensation, Contact, and Description of Satisfaction Value. In this study, the results obtained from several input attributes that create a causal relationship when classifying satisfied and dissatisfied consumers. The results obtained can provide benefits for the Shopee company in increasing customer satisfaction. After carrying out the testing process, it can be concluded that the Naive Bayes Algorithm is an algorithm that is suitable for use in the classification process for measuring Shopee Indonesia's consumer satisfaction level among the Millennial Generation, with an accuracy rate of 89.65%.

Keywords: *Classification, Consumer Satisfaction, Data Mining, E-Commerce, Naïve Bayes Algorithm.*

TEKNIK DATA MINING MENGGUNAKAN ALGORITMA NAIVE BAYES UNTUK MEMREDIKSI KEPUASAN KONSUMEN SHOPEE DIKALANGAN GENERASI MILENIAL

Abstrak

Shopee merupakan salah satu *e-commerce* terbesar yang saat ini sedang marak digunakan oleh Generasi Milenial. Penggunaan Shopee sendiri sangat mempermudah konsumennya dalam melakukan proses transaksi. Shopee sendiri memiliki komitmen dalam menjaga dan meningkatkan kepuasan konsumennya agar tidak berpindah ke kompetitor yang lain. Namun tidak dapat dipungkiri masih banyak sekali kasus-kasus yang dapat merugikan konsumen pada saat menggunakan *platform* tersebut. Dengan adanya kasus-kasus yang terjadi, sangat memungkinkan adanya pengaruh besar terhadap tingkat kepuasan konsumen pada *platform* tersebut. Konsumen akan merasa puas ketika produk atau jasa yang digunakan bisa memenuhi ekspektasi konsumen. Penelitian ini dibuat dengan tujuan untuk memprediksi tingkat kepuasan konsumen Shopee Indonesia di kalangan Generasi Milenial. Penelitian ini melakukan penerapan *data mining* dengan menggunakan Algoritma Naive Bayes. Algoritma Naive Bayes sendiri adalah suatu klasifikasi kemungkinan sederhana yang dapat menghitung seluruh kemungkinan dengan menggabungkan sejumlah kombinasi dan frekuensi suatu nilai dari basis data yang didapatkan. Atribut yang digunakan dalam melakukan penelitian ini mencakup Nama, Jenis Kelamin, Umur, *Price, Performance* dan *Efficiency, Fulfillment, Reliability, Control and Security, Responsiveness, Compensation, Contact*, dan Keterangan Nilai Kepuasan. Pada penelitian ini didapatkan hasil yang diperoleh dari beberapa atribut masukan yang menciptakan hubungan sebab akibat ketika mengklasifikasikan konsumen yang puas dan tidak puas. Hasil yang telah diperoleh tersebut dapat memberikan manfaat bagi perusahaan Shopee dalam meningkatkan kepuasan konsumennya. Setelah melakukan proses pengujian dapat disimpulkan bahwa Algoritma Naive Bayes adalah algoritma yang cocok digunakan dalam melakukan proses klasifikasi terhadap pengukuran tingkat kepuasan konsumen Shopee Indonesia di kalangan Generasi Milenial, dengan menghasilkan tingkat akurasi sebesar 89.65%.

Kata kunci: *Algoritma Nive Bayes, E-Commerce, Data Mining, Kepuasan Konsumen, Klasifikasi.*

1. PENDAHULUAN

Di era digital seperti ini semuanya serba cepat dan praktis terutama bagi kaum generasi milenial yang sangat senang sekali berbelanja, generasi milenial membutuhkan suatu *platform* belanja yang sangat mudah tanpa harus mengeluarkan tenaga untuk datang langsung ke tempat yang menjual produk yang diinginkan. Menurut penelitian terbaru, diungkapkan bahwa generasi milenial terutama di Indonesia merupakan penyumbang terbesar dalam menggunakan *e-commerce* sebagai wadah untuk berbelanja [1]. Dapat kita ketahui *e-commerce* sendiri adalah tempat atau wadah di dalam melakukan transaksi jual beli tanpa harus bertemu secara langsung, melainkan dapat dilakukan secara elektronik melalui media internet [2]. Dapat dilihat *e-commerce* membawa peluang bisnis yang sangat besar, sehingga banyak sekali perusahaan yang berlomba-lomba menyediakan layanan secara online baik sebagai pengganti atau sebagai tambahan untuk layanan *offline* produk mereka. Akibat dari pertumbuhan *e-commerce* yang terus meningkat membuat Shopee ikut terjun meramaikan industri ini. Shopee sendiri merupakan salah satu e-marketplace yang meningkat pesat di Indonesia, karena Shopee sudah berhasil menarik minat konsumen agar lebih tertarik bertransaksi melalui situs tersebut sebagai wadah jual beli secara *online* [3]. Saat pertama kali Shopee masuk ke Indonesia, shopee langsung memberikan penawaran menarik yaitu promo gratis ongkir kemana saja dengan minimal pembelian. Shopee juga memiliki berbagai macam promo menarik mulai dari *flash sale*, *cashback*, promo gratis ongkir, serta promo besar-besaran lainnya yang diadakan oleh berbagai brand [4]. Pelayanan yang diterima oleh konsumen dapat mempengaruhi tingkat kepuasan mereka dalam berbelanja. Pelayanan yang bagus dapat mendorong konsumen untung berkomitmen terhadap produk serta pelayanan perusahaan. Kepuasan yang diperoleh konsumen saat berbelanja merupakan perasaan senang atau kecewa dari hasil membandingkan kinerja suatu produk dengan layanan yang diperoleh serta ekspektasinya (Kotler & Keller, 2016). Jika pelayanan yang diterima oleh konsumen tidak sesuai ekspektasi tentu konsumen merasa tidak puas, sebaliknya jika pelayanan yang diberikan sesuai ekspektasi, maka konsumen akan merasa puas, dan jika pelayanan tersebut melebihi ekspektasi konsumen, maka konsumen akan sangat puas [5]. Konsumen cenderung akan membandingkan pelayanan yang konsumen terima dengan perusahaan yang lain sebagai tolak ukur tingkat kepuasan.

Konsumen adalah orang yang berhak memberikan kritik maupun saran terhadap kualitas pelayanan yang konsumen dapatkan dari sebuah

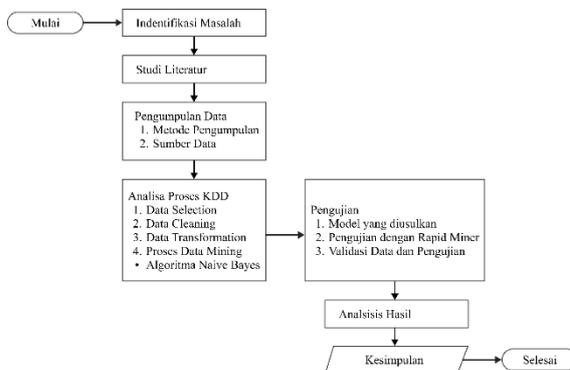
produk maupun perusahaan. Penilaian kepuasan konsumen pasti berdasarkan perbandingan dari apa yang konsumen harapkan dengan yang diterima. Ketidakpuasan yang dialami konsumen ini dapat menjadi faktor utama yang bisa menyebabkan konsumen berpindah merek. Menurut (Dharmmesta, 2002) seseorang konsumen yang mengalami ketidakpuasan mempunyai kemungkinan akan merubah perilaku keputusan belinya dengan mencari alternatif merek lain pada konsumsi berikutnya untuk meningkatkan kepuasannya [6]. Kepuasan konsumen sendiri dapat diartikan sebagai suatu keadaan dimana saat konsumen menerima produk yang konsumen pesan dan produk tersebut memenuhi harapan dari konsumen. Saat konsumen menerima produk yang sesuai maka konsumen akan merasa puas sehingga konsumen akan loyal terhadap perusahaan dan akan menggunakan jasa perusahaan lagi, sebaliknya jika produk yang diterima jauh dari harapan konsumen, maka konsumen akan merasa kecewa. Harapan konsumen mengenai produk yang konsumen pilih pada suatu *marketplace* dapat diketahui melalui pengalaman pribadi konsumen, informasi dari orang lain, serta dari iklan atau promosi yang pernah didengar dan dilihat oleh konsumen [3]. Karena ketidakpuasan yang dialami oleh konsumen sehingga kecenderungan menimbulkan sikap negatif terhadap merek, dan penyedia jasa dimana konsumen berbelanja, kemungkinan terbesar yang terjadi adalah konsumen tidak akan melakukan pembelian ulang (*repurchase*), berganti dengan merek yang lain (*brand switching*), melakukan komplain, dan bahkan beralih ke marketplace yang lain [7].

Melihat pada saat ini banyak sekali Generasi Milenial yang sering berbelanja menggunakan *platform* Shopee. Penelitian sekarang ini menggunakan Naïve Bayes sebagai algoritma yang dipilih untuk melakukan klasifikasi terhadap tingkat kepuasan konsumen saat berbelanja di Shopee. Dapat diketahui Algoritma Naïve Bayes sendiri adalah suatu klasifikasi kemungkinan sederhana yang dapat menghitung seluruh kemungkinan dengan menggabungkan sejumlah kombinasi dan frekuensi suatu nilai dari basis data yang didapatkan [8]. Algoritma Naïve Bayes ini memanfaatkan teorema bayes dan memperkirakan seluruh atribut yang bebas dan saling lepas yang dapat diberikan oleh suatu nilai pada kelas variable. Klasifikasi Naïve Bayes sendiri dapat melakukan prediksi di masa yang akan datang berdasarkan dari pengalaman di masa sebelumnya [9]. Untuk melakukan evaluasi apakah perusahaan telah memberikan kualitas pelayanan yang baik terhadap konsumen perlu adanya evaluasi yang dilakukan oleh konsumen untuk mengetahui seberapa puas konsumen saat berbelanja pada e-marketplace ini. Dari permasalahan yang telah dijelaskan, maka

penulis melakukan penelitian dengan judul "Teknik Data Mining menggunakan Algoritma Naïve Bayes untuk Memprediksi Kepuasan Konsumen Shopee Indonesia di Kalangan Generasi Milenial". Penelitian ini akan menghasilkan klasifikasi dari tingkat kepuasan konsumen Shopee di Indonesia, dengan melakukan pendekatan data mining berdasarkan hasil dari klasifikasi menggunakan Algoritma Naïve Bayes [10].

2. METODE PENELITIAN

Tahapan penelitian ini terdiri dari metodologi penelitian, dimana metodologi penelitian sendiri adalah proses atau alur bagaimana cara kita memperoleh data dan informasi sampai dengan mengeksekusi data dengan tujuan mencari data yang akurat, yang akan membantu dalam penelitian tugas akhir. Alur pada Metodologi Penelitian dapat dilihat pada Gambar 1.



Gambar 1. Tahapan Penelitian

2.1. Identifikasi Masalah

Semua penelitian akan selalu didahului dengan identifikasi masalah, hal ini berguna untuk peneliti agar fokus pada titik permasalahan. Berdasarkan latar belakang yang sudah dijelaskan sebelumnya, ada beberapa garis besar permasalahan yang dapat disimpulkan yaitu:

1. Bagaimana cara memprediksi kepuasan konsumen Shopee Indonesia serta memberikan penangan yang lebih tepat?
2. Bagaimana Algoritma Naive Bayes dapat membantu proses klasifikasi untuk memprediksi kepuasan konsumen Shopee Indonesia dan melihat tingkat keakuratan algoritma Naive Bayes?
3. Apakah Algoritma Naive Bayes sendiri dapat membantu pihak perusahaan dalam melihat prediksi tingkat kepuasan konsumen Shopee Indonesia dikalangan Generasi Milenial?

2.2. Studi Literatur

Tahap ini dilakukan dengan cara melakukan pengumpulan informasi yang dibutuhkan dan berhubungan dengan penelitian yang dikerjakan oleh penulis yang diambil dari berbagai macam media.

Penulis mengumpulkan data-data dengan cara mempelajari beberapa jurnal, buku maupun referensi lain yang terkait dengan penelitian yang menggunakan Algoritma Naive Bayes yang dapat digunakan sebagai acuan yang baik dan relevan.

2.3. Pengumpulan Data

Setelah melakukan proses pada studi literatur, penulis melakukan pengumpulan data yang ditujukan kepada Generasi Milenial sebagai responden utama.

2.3.1. Metode Pengumpulan Data

Data dalam penelitian tugas akhir ini didapatkan dari hasil survey kuesioner secara *online* menggunakan *Google Form* yang diberikan kepada Generasi Milenial yang menggunakan Shopee sebanyak 657 responden. Adapun skala pengukuran yang dipakai dalam kuesioner ini menggunakan Skala Likert.

2.3.2. Sumber Data

Sumber data yang digunakan dalam penelitian ini yaitu Data Primer dimana data primer sendiri adalah data yang dicari, diamati, dan diperoleh secara langsung oleh penulis yang melakukan suatu penelitian. Data penelitian ini sendiri adalah data dari hasil kuesioner yang telah disebarluaskan kemudian diisi oleh beberapa Generasi Milenial yang sudah pernah menggunakan *e-commerce* Shopee.

2.4. Analisis Data Training dan Data Testing

Dalam memprediksi tingkat kepuasan konsumen Shopee Indonesia dikalangan Generasi Milenial penulis menggunakan Algoritma Naive Bayes dalam melakukan proses klasifikasi, sehingga sebelum melakukan proses pada data tersebut, penulis harus menganalisa data yang akan digunakan. Disini data yang digunakan adalah sampel dari hasil kuesioner yang telah dibagikan kepada beberapa orang. Dalam kuesioner yang telah dibagikan tersebut terdapat dua belas atribut yang belum melewati proses *cleaning data*. Dimana dari dua belas atribut yang ada akan dibagi lagi menjadi atribut *output* dan atribut target. Berdasarkan kuesioner yang telah disebarluaskan, tercatat 657 responden yang telah mengisi kuesioner tersebut. Atribut yang ada berjumlah dua belas atribut, atribut tersebut akan melakukan proses *preparation data*, dimana *preparation data* adalah proses dimana menentukan data mana saja yang berkualitas yang harus digunakan.

2.5. Analisis Proses Knowledge Discovery in Database

Istilah yang paling sering digunakan pada saat melakukan pengolahan data yang besar adalah *data mining* dan *Knowledge Discovery in Database* (KDD). Proses KDD sendiri digunakan untuk

mendapatkan pengetahuan dari *database* yang ada, dimana dalam suatu *database* terdapat berbagai table-tabel yang saling berhubungan antara satu dengan yang lain [14]. Proses KDD dan *Data Mining* sering digunakan secara bergantian, walaupun mempunyai konsep yang berbeda tetapi proses kdd dan *data mining* saling berkaitan. Adapun tahapan dari proses KDD, yaitu:

2.5.1. Seleksi Data (Data Selection)

Selection Data adalah proses dalam memilih data yang akan digunakan kemudian membagi data menjadi *data training* dan *data testing*, kemudian data yang digunakan pada penelitian ini adalah data primer yang berasal dari sampel kuesioner. Pada proses seleksi data sendiri data kuantitatif yang digunakan akan diminimalkan tetapi tetap mempresentasikan data aslinya. Data pada *database* tidak semuanya dipakai tetapi hanya data yang sesuai untuk melakukan analisis dalam melakukan penelitian.

2.5.2. Pembersihan Data (Data Cleaning)

Tahap kedua pada proses KDD adalah *Cleaning Data*. Dimana *Cleaning Data* adalah proses dimana menghapus beberapa atribut yang sekiranya tidak diperlukan dalam melakukan klasifikasi data. Pembersihan data yang dilakukan guna untuk menghilangkan *noise* dan data yang tidak konsisten atau data yang tidak relevan, tidak *missing value* dan tidak *redundant*.

- a) Dalam Penelitian yang dilakukan oleh penulis, atribut yang akan digunakan dalam proses klasifikasi adalah Nama, Jenis Kelamin, Usia, *Price*, *Performance and Efficiency*, *Fulfillment*, *Reliability*, *Control and Security*, *Responsiveness*, *Compensation*, *Contact*, dan Keterangan Nilai Kepuasan.
- b) Data yang didapatkan dari hasil kuesioner masih terdapat beberapa *missing value*, dimana data yang *missing value* tersebut akan diganti untuk membantu proses klasifikasi.

masuk kedalam aplikasi dengan mudah. *Fulfillment* mengukur kemampuan Shopee dalam memenuhi penyediaan layanan secara tepat waktu dan akurat yang dibutuhkan konsumen. *Reliability* adalah dimensi yang mengukur kemampuan Shopee untuk melakukan layanan yang sesuai dengan apa yang dijanjikan secara akurat dan handal atau terpercaya. *Control and Security* mengukur kemampuan Shopee dalam menjamin keamanan data dan transaksi konsumen dengan aman dan tanpa membocorkannya atau menyalahgunakan data tersebut. *Responsiveness* mengukur kemampuan Shopee untuk memberikan informasi yang akurat kepada konsumen ketika timbul suatu masalah dan memiliki mekanisme pemberian garansi dari masalah tersebut. *Compensation* mengukur kemampuan Shopee dalam memberikan kompensasi kepada konsumen jika terjadi kesalahan atau kegagalan sistem. *Contact* adalah dimensi yang mengukur kemampuan Shopee dalam memberikan informasi terkait produk dan layanan, perusahaan dapat menyediakan

6 *Fulfillment* Digunakan Atribut

7 *Reliability* Digunakan Atribut

8 *Control and Security* Digunakan Atribut

9 *Responsiveness* Digunakan Atribut

10 *Compensation* Digunakan Atribut

11 *Contact* Digunakan Atribut

Tabel 1. Atribut Data Penelitian

No.	Atribut	Proses	Penjelasan
1	Nama	Digunakan Atribut	Nama Generasi Milenial
2	Jenis Kelamin	Digunakan Atribut	Laki-Perempuan
3	Usia	Digunakan Atribut	Umur Generasi Milenial
4	<i>Price</i>	Digunakan Atribut	Harga yang ditawarkan oleh Shopee
5	<i>Performance and Efficiency</i>	Digunakan Atribut	<i>Performance & Efficiency</i> mengukur kemampuan Shopee dalam memberikan akses kemudahan dalam mencari informasi yang dibutuhkan dan akses keluar-

12	Keterangan Nilai Kepuasan	Digunakan sebagai Label	komunikasi yang nyaman antara pelanggan dan karyawan secara online atau melalui telepon. Tingkat kepuasan Puas/Tidak Puas menggunakan Shopee
----	---------------------------	-------------------------	---

2.5.3. Transformasi Data (*Data Transformation*)

Proses *transformation data* adalah dimana sebelumnya data yang telah melewati proses *preprocessing* kemudian data tersebut akan disimpan kedalam bentuk data yang sudah siap digunakan pada tools yang akan digunakan nanti. Data akan diubah atau digabung ke dalam format yang sesuai untuk diproses dalam *data mining*.

Tabel 2. Klasifikasi Pada Atribut

No.	Atribut	Klas
1	Nama	Nama Generasi Milenial
2	Jenis Kelamin	Laki-Laki
3		Perempuan
3	Umur	Umur Generasi Milenial
4	Price	Sangat Setuju
4		Setuju
4		Tidak Setuju
5	Performance and Efficiency	Sangat Tidak Setuju
5		Sangat Setuju
5		Setuju
6	Fulfillment	Tidak Setuju
6		Sangat Tidak Setuju
6		Sangat Setuju
7	Reliability	Setuju
7		Tidak Setuju
7		Sangat Tidak Setuju
8	Control and Security	Sangat Setuju
8		Setuju
8		Tidak Setuju
9	Responsiveness	Sangat Tidak Setuju
9		Sangat Setuju
9		Setuju
10	Compensation	Tidak Setuju
10		Sangat Tidak Setuju
10		Sangat Setuju
11	Contact	Setuju
11		Tidak Setuju
11		Sangat Tidak Setuju
12	Keterangan Nilai Kepuasan	Puas
12		Tidak Puas

Pada Tabel 2 menjelaskan dua belas atribut yang digunakan sebagai *input* yang sudah memiliki klasifikasi masing-masing yang akan mempermudah dalam proses pendekatan *data mining*, kemudian dua belas atribut yang sudah ada tersebut akan dianalisa menggunakan algoritma Naive Bayes.

2.6. Metode yang Digunakan

Metode yang akan digunakan penulis untuk melakukan klasifikasi terhadap tingkat kepuasan kosnumen Shopee Indonesia adalah dengan menggunakan metode naive bayes.

2.6.1. Algoritma Naive Bayes Classifier

Naive Bayes adalah metode klasifikasi yang prosesnya sangat sederhana, metode ini sendiri berakar dari teorema Bayes dan hipotesis kebebasan. Klasifikasi pada naive bayes menggunakan metode probabilitas dan statistik dengan menggunakan probabilitistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset. Dibandingkan dengan algoritma lainnya Algoritma Naive Bayes adalah algoritma yang paling banyak digunakan karena tingkat efisiensi dan kemampuannya yang besar dalam menggabungkan bukti. Algoritma Naive Bayes ini memberikan suatu prediksi di masa depan menurut pengalaman yang sudah pernah terjadi sebelumnya. Keunggulan utama dari Algoritma Naive Bayes adalah tidak memerlukan data yang besar dalam melakukan proses klasifikasi, tetapi tingkat akurasi yang dihasilkan tinggi dan data diproses dengan cepat. Untuk menghitung setiap kelas keputusan pada naive bayes harus memiliki syarat bahwa kelas keputusan adalah benar (Olson, 2008). Menurut teori Thomas bayes formulasi umum yang dikemukakan adalah sebagai berikut:

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

Naive Bayesian Classifier mengasumsikan dengan kuat maksud dari independensi. Independensi yang kuat dalam atribut adalah atribut yang dihasilkan adalah atribut yang bebas, dimana data tidak berkaitan dengan ada atau tidaknya atribut yang lain dalam data yang sama karena tidak adanya keterkaitan antar atribut. Berdasarkan asumsi yang telah dijelaskan, berlaku juga persamaan sebagai berikut:

$$P(H|X) = P(H) \sum_{i=1}^n P(X_i|H) \quad (2)$$

Keterangan :

X : Data yang kelasnya belum diketahui

H : Hipotesis dari data X yang nilainya lebih spesifik

P(H|X) : Probabilitas hipotesis H berdasarkan dari kondisi X

P(X|H) : Probabilitas hipotesis X berdasarkan dari kondisi H

P(H) : Probabilitas hipotesis H

P(X) : Probabilitas hipotesis X

Ide dasar dari aturan bayes adalah bahwa hasil dari hipotesis atau peristiwa (H) dapat diperkirakan berdasarkan pada beberapa bukti (X) yang diamati. Beberapa hal yang perlu diperhatikan :

- a. Sebuah probabilitas awal atau prior H atau $P(H)$ adalah probabilitas dari suatu hipotesis sebelum bukti diamati.
- b. Sebuah probabilitas akhir H atau $P(H|X)$ adalah probabilitas dari suatu hipotesis setelah bukti diamati [15].

2.7. Pengujian Dengan RapidMiner

RapidMiner merupakan perangkat lunak yang bersifat *open source* yang digunakan sebagai alat pembelajaran pada *data mining*. RapidMiner merupakan solusi yang sangat dibutuhkan untuk membantu melakukan analisis terhadap *data mining*, *text mining* dan analisis prediksi. Penelitian ini akan mengklasifikasi data dengan melakukan pendekatan *data mining*, dimana data tersebut akan diproses menggunakan *tools* yang sudah dipilih yaitu RapidMiner. *Tools* RapidMiner sendiri sudah menyediakan komponen yang lengkap dalam membantu proses *data mining*. *Tools* RapidMiner sendiri memiliki beberapa sifat sebagai berikut:

- a. *Tools* RapidMiner sendiri ditulis menggunakan bahasa pemrograman Java yang bisa digunakan disemua *system* operasi.
- b. Terjadi penemuan yang kemudian menjadi *operation trees*.
- c. Adanya representasi XML agar bias melakukan proses pertukaran data.
- d. *Tools* RapidMiner juga menerapkan *multi-layer* yang dapat memastikan tampilan akan lebih efisien.
- e. Memiliki GUI, *command line mode*, dan Java API yang dapat dipanggil dari program lain [17].

2.8. Evaluasi Validasi Perhitungan Akurasi dan Confusion Matrix

Proses validasi ini akan melewati beberapa kali percobaan dimana pada setiap percobaan memiliki komposisi data yang berbeda-beda. Proses validasi yang dilakukan menggunakan *Confusion Matrix* ini bertujuan untuk melihat tingkat akurasi, *recall*, dan presisi pada data, dimana akurasi sendiri adalah persentase klasifikasi pada pengujian data set yang benar. Proses presisi sendiri adalah persentase hasil klasifikasi dimana, model mana yang baik. Dan terakhir adalah *recall*, dimana *recall* sendiri adalah pengukuran tingkat positif sebenarnya. Tahap ini adalah tahap dimana pengujian data menggunakan *Confusion Matrix* kemudian mencari tingkat akurasi yang terbaik berdasarkan model yang digunakan.

Tabel 3. Confusion Matrix

Classification Naïve Bayes	Layak(+)	Tidak Layak (-)
Layak (+)	TN	FN
Tidak Layak (-)	FP	TP

Dari Tabel 4 dapat hitung akurasi sebagai berikut. Akurasi = $x \cdot 100\%$ [18].

3. HASIL DAN PEMBAHASAN

Penelitian ini menerapkan pendekatan implementasi *data mining* dengan menggunakan algoritma Naïve Bayes, dimana data yang digunakan adalah data yang didapatkan dari hasil kuesioner online yang telah disebar oleh penulis, dan mendapatkan 657 responden. *Tools* yang dipilih penulis dalam membantu proses klasifikasi adalah RapidMiner, dimana aplikasi RapidMiner sendiri sangat membantu dalam proses ekstraksi data dengan metode *data mining*.

3.1. Alur Proses Penelitian

Data dalam penelitian tugas akhir ini didapatkan dari hasil *survey* yang dilakukan dengan cara membagikan kuesioner *online* yang dilakukan oleh penulis kepada Generasi Milenial di Indonesia yang menggunakan Shopee. Metode yang digunakan dalam penelitian ini yaitu model algoritma Naive Bayes. Berikut alur penelitian:

- a. Penulis akan membuat kuesioner *online*, dengan tujuan mencapai target sebanyak- banyaknya, kemudian data yang didapatkan dari kuesioner *online* tersebut menjadi data awal untuk mengetahui dan mengukur tingkat kepuasan konsumen Shopee Indonesia.
- b. Melakukan perhitungan terhadap data kuesioner *online* yang sudah masuk.
- c. Menerapkan *Data Selection*, untuk memilih himpunan data yang akan digunakan yaitu data kuisisioner *online* konsumen Shopee Indonesia.
- d. *Cleaning data* adalah proses dimana menghapus beberapa atribut yang sekiranya tidak diperlukan dalam melakukan klasifikasi data atau yang tidak memiliki nilai atribut lengkap (*missing value*).
- e. Hasil perhitungan puas dan tidak puas kemudian diolah menggunakan rumus metode Naïve Bayes.
- f. Proses validasi yang dilakukan menggunakan *Confusion Matrix* ini bertujuan untuk melihat tingkat akurasi, *recall*, dan presisi data.
- g. *Tools* yang akan digunakan oleh peneliti adalah RapidMiner, karena RapidMiner sendiri memiliki komponen yang lengkap untuk melakukan proses *data mining*.

3.2. Proses Pengolahan Data

Data yang digunakan oleh penulis adalah data hasil kuesioner *online*, dimana data yang diperoleh berjumlah 657 responden. Atribut yang digunakan penulis dalam melakukan penelitian ini berjumlah dua belas, yang terdiri dari Nama, Jenis Kelamin, Umur, *Price*, *Performance and Efficiency*, *Fulfillment*, *Reliability*, *Control and Security*, *Responsiveness*, *Compensation*, *Contact*, dan Keterangan Nilai Kepuasan. Dari 657 responden yang telah mengisi kuesioner terdapat 641 konsumen Shopee yang merasa puas dan 16 konsumen Shopee

yang merasa tidak puas. Berikut adalah data dari hasil kuesioner *online* yang telah diisi oleh konsumen Shopee di Indonesia:

Gambar 2. Data Set Kuesioner Konsumen Shopee

Sebelum digunakan untuk membantu proses klasifikasi, data tersebut harus melewati proses *cleaning data* dimana pada RapidMiner sendiri terdapat operator yang dapat membantu dalam pembersihan data yaitu *Replace Missing Value*. Berikut adalah data yang sudah melewati proses pengolahan data awal, dimana pada bagian *Missing* nilai yang diberikan adalah 0, yang artinya tidak terdapat data yang *missing value*.

Name	Type	Missing	Statistika	Filter (26 / 30 atribut)	Search Bar (Atribut)
Bagaimana kepuasan anda tentang...	Nominal	0	Puas Tidak Puas	Value Puas (641), Tidak Puas (16)	
Nama Responden	Nominal	0	Laki-laki Perempuan	Value A (3), Anisa (5), [025 more]	
Domisili (Kabupaten/Kota)	Nominal	0	Yogyakarta (1)	Value Ketapang (60), Salatiga (36), [27	
Jenis Kelamin	Binomial	0	Laki-Laki Perempuan	Value Perempuan (600), Laki-Laki (157)	
Usia Responden	Integer	0	15 38	Value Rangga (20), 828	
Harga yang ditawarkan di Shop...	Nominal	0	Sangat Tidak Setuju (5) Tidak Setuju (372) Setuju (372) Sangat Setuju (269)	Value Setuju (372), Sangat Setuju (269)	
Produk yang dibutuhkan konsum...	Nominal	0	Tidak Setuju (14) Tidak Setuju (353) Setuju (353) Sangat Setuju (200)	Value Setuju (353), Sangat Setuju (200)	
Informasi produk di Shopee m...	Nominal	0	Sangat Tidak Setuju (1) Tidak Setuju (373) Setuju (373) Sangat Setuju (262)	Value Setuju (373), Sangat Setuju (262)	
Shopee mudah untuk diakses ...	Nominal	0	Tidak Setuju (14) Tidak Setuju (370) Setuju (370) Sangat Setuju (273)	Value Setuju (370), Sangat Setuju (273)	
Aplikasi shopee merespon su...	Nominal	0	Luar Setuju (423) Tidak Setuju (423) Sangat Setuju (145)	Value Setuju (423), Sangat Setuju (145)	
Pilihan metode pembayaran su...	Nominal	0	Tidak Setuju (15) Tidak Setuju (396) Setuju (396) Sangat Setuju (246)	Value Setuju (396), Sangat Setuju (246)	
Untuk memuat halaman Shope...	Nominal	0	Sangat Tidak Setuju (12) Tidak Setuju (356) Setuju (356) Sangat Setuju (148)	Value Setuju (356), Sangat Setuju (148)	
Tersedia fasilitas pencarian un...	Nominal	0	Tidak Setuju (5) Tidak Setuju (369) Setuju (369) Sangat Setuju (283)	Value Setuju (369), Sangat Setuju (283)	
Shopee selalu memberikan gra...	Nominal	0	Sangat Tidak Setuju (7) Tidak Setuju (378) Setuju (378) Sangat Setuju (204)	Value Setuju (378), Sangat Setuju (204)	
Seting memberikan potongan ...	Nominal	0	Sangat Tidak Setuju (4) Tidak Setuju (383) Setuju (383) Sangat Setuju (195)	Value Setuju (383), Sangat Setuju (195)	
Harga promo yang ditawarkan ...	Nominal	0	Sangat Tidak Setuju (9) Tidak Setuju (371) Setuju (371) Sangat Setuju (237)	Value Setuju (371), Sangat Setuju (237)	
Barang yang dikirim Shopee se...	Nominal	0	Sangat Tidak Setuju (1) Tidak Setuju (496) Setuju (496) Sangat Setuju (107)	Value Setuju (496), Sangat Setuju (107)	
Sistem Shopee bekerja deng...	Nominal	0	Sangat Tidak Setuju (3) Tidak Setuju (400) Setuju (400) Sangat Setuju (175)	Value Setuju (400), Sangat Setuju (175)	
Mudah melakukan transaksi di ...	Nominal	0	Sangat Tidak Setuju (2) Tidak Setuju (362) Setuju (362) Sangat Setuju (281)	Value Setuju (362), Sangat Setuju (281)	
Customer service shopee dap...	Nominal	0	Sangat Tidak Setuju (6) Tidak Setuju (464) Setuju (464) Sangat Setuju (145)	Value Setuju (464), Sangat Setuju (145)	
Merasa aman mengenai data p...	Nominal	0	Sangat Tidak Setuju (6) Tidak Setuju (409) Setuju (409) Sangat Setuju (165)	Value Setuju (409), Sangat Setuju (165)	
Mudah menghapus akun peng...	Nominal	0	Sangat Tidak Setuju (5) Tidak Setuju (452) Setuju (452) Sangat Setuju (109)	Value Setuju (452), Sangat Setuju (109)	
Customer service Shopee cep...	Nominal	0	Sangat Tidak Setuju (4) Tidak Setuju (436) Setuju (436) Sangat Setuju (116)	Value Setuju (436), Sangat Setuju (116)	
Shopee memberikan pesan inf...	Nominal	0	Sangat Tidak Setuju (6) Tidak Setuju (392) Setuju (392) Sangat Setuju (195)	Value Setuju (392), Sangat Setuju (195)	
Shopee selalu memberikan not...	Nominal	0	Sangat Tidak Setuju (2) Tidak Setuju (361) Setuju (361) Sangat Setuju (238)	Value Setuju (361), Sangat Setuju (238)	
Shopee memberikan garansi 1...	Nominal	0	Sangat Tidak Setuju (2) Tidak Setuju (394) Setuju (394) Sangat Setuju (201)	Value Setuju (394), Sangat Setuju (201)	
Shopee menawarkan retur bar...	Nominal	0	Sangat Tidak Setuju (6) Tidak Setuju (413) Setuju (413) Sangat Setuju (181)	Value Setuju (413), Sangat Setuju (181)	
Shopee memberikan kompens...	Nominal	0	Sangat Tidak Setuju (11) Tidak Setuju (382) Setuju (382) Sangat Setuju (139)	Value Setuju (382), Sangat Setuju (139)	
Layanan komunikasi dengan p...	Nominal	0	Tidak Setuju (9) Tidak Setuju (378) Setuju (378) Sangat Setuju (270)	Value Setuju (378), Sangat Setuju (270)	

Gambar 3. Hasil Pengolahan Data Awal

3.3. Proses Data Mining Pada RapidMiner

Setelah melakukan proses pengolahan data diawal, data akan diproses menggunakan *tools* yang sudah ditetapkan oleh penulis yaitu Rapid Miner, kemudian pengujiannya akan menggunakan operator yang sudah tersedia pada *tools* RapidMiner yaitu *Cross Validation*. Berikut adalah data yang akan diimport ke dalam RapidMiner:

Hasil	Nama	Jenis Kelamin	Usia	Domisili	Harga yang ditawarkan di Shopee relatif lebih terjangkau	Produk yang dibutuhkan konsumen dapat dengan mudah ditemukan di Shopee	Shopee mudah untuk diakses pengguna
Binomial	Polynomial	Binomial	Polynomial	Polynomial	Polynomial	Polynomial	Polynomial
Label	Id	Atribut	Atribut	Atribut	Atribut	Atribut	Atribut
Puas	Valentino Kristian	Laki-Laki	21	Salatiga	Sangat Setuju	Sangat Setuju	Sangat Setuju
Puas	Sangaya Agustinus	Laki-Laki	21	Yogyakarta	Setuju	Setuju	Sangat Setuju
Tidak Puas	Francisca Osamu Pedorova Berlin	Perempuan	19	Ketapang	Setuju	Setuju	Tidak Setuju
Puas	Sarah Wandita Napinupu	Perempuan	21	Ketapang	Sangat Setuju	Sangat Setuju	Tidak Setuju

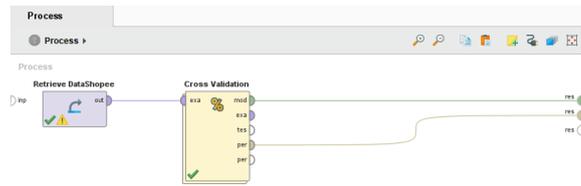
Gambar 3. Data yang Diimport ke RapidMiner

Keterangan: Nama Variabel ■
 Tipe Data ■
 Label/Atribut ■

3.4. Pengujian Naive Bayes Dengan RapidMiner

3.4.1. Pengujian Probabilitas

Proses pengujian probabilitas adalah melakukan import data *training* sebanyak 657 data responden. Berikut adalah pemodelan validasi:



Gambar 4. Pemodelan Data Training

Setelah melakukan pemodelan pada pengujian probabilitas, tahap selanjutnya adalah *simple distribution* dimana pada tahap ini merupakan proses *import data* yang digunakan untuk menentukan nilai dari data kelas puas dan data kelas yang tidak puas. Berikut pengujian *probabilitas prior* dapat dilihat pada Gambar 5:

SimpleDistribution
Distribution model for label attribute Bagaimana kepuasan anda tentang Price, Performance & Efficiency, Fulfillment, Reliability, Control and Security, Responsiveness, Compensation, Contact yang diberikan oleh Shopee?
Class Puas (0.976) 29 distributions
Class Tidak Puas (0.024) 29 distributions

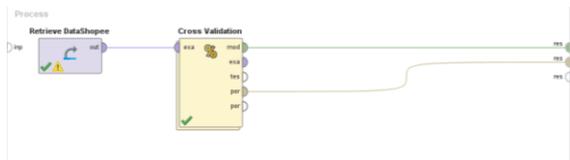
Gambar 5. Pengujian Probabilitas Prior

Berdasarkan Gambar 5 dapat disimpulkan bahwa Kelas Puas memiliki nilai probabilitas sebesar

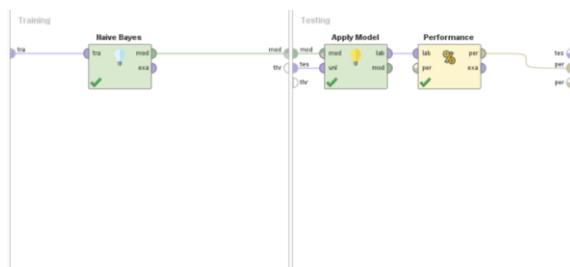
0.976 dan Kelas Tidak Puas memiliki nilai probabilitas sebesar 0.024.

3.4.2. Pemodelan Proses Data Testing

Pada pemodelan proses *data testing* dilakukan dengan cara membuat struktur model, dimana proses ini akan membantu dalam mencari tingkat akurasi yang dihasilkan dan ketepatan prediksi.



Gambar 6. Desain Model Pengujian



Gambar 7. Pemodelan Naive Bayes pada RapidMiner saat pengujian

Dapat dilihat pada Gambar 6 adalah struktur model yang diterapkan pada RapidMiner. Bagian *Retrieve Data Shopee* adalah *input* dari data target yang akan diklasifikasi, kemudian menerapkan komponen Algoritma Naive Bayes sebagai metode yang digunakan dalam proses klasifikasi.

3.4.3. Hasil Pengujian Algoritma Naive Bayes

Berikut adalah hasil *confusion matrix* menggunakan Algoritma Naive Bayes yang didapatkan dari hasil pengujian berdasarkan *data training* yang diambil dari data kuesioner online yang berjumlah 657 responden, dari *table 5* dapat dilihat bahwa Algoritma Naive Bayes mendapatkan hasil akurasi sebesar 89.65%, dengan *Class Precision* sebesar 98.97%, dan *Recall* sebesar 90.33%. Dari hasil klasifikasi menggunakan Algoritma Naive Bayes dapat dihasilkan bentuk *confusion matrix* sebagai berikut.

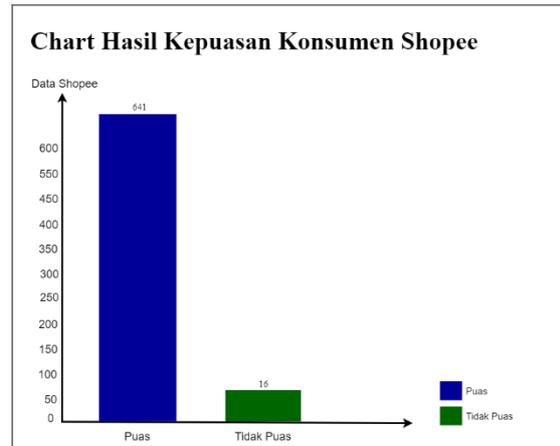
Tabel 4. Hasil Pengujian Algoritma Naive Bayes

	<i>true</i> Puas	<i>true</i> Tidak Puas	<i>Class</i> <i>Precision</i>
pred.Puas	579	6	98.97%
pred.Tidak Puas	62	10	13.89%
class recall	90.33%	62.50%	

3.4.4. Grafik Klasifikasi

Berdasarkan jumlah responden sebanyak 657 responden, terdapat 641 responden yang puas dan 16 responden yang tidak puas. Berikut adalah hasil implementasi berupa *Chart* menggunakan Algoritma

Naive Bayes berdasarkan hasil kepuasan konsumen Shopee di Indonesia.



Gambar 8. Chart Hasil Kepuasan Konsumen Shopee

Berdasarkan hasil yang telah di dapatkan, akan menghasilkan sebuah probabilitas kemunculan dari setiap nilai dari suatu atribut. Berikut merupakan hasil probabilitas hasil yang telah diproses.

Attribute	Parameter	Puas	Tidak Puas
Nama Responden	value=Valentino Kristian Reynald	0.002	0.000
Nama Responden	value=Yehochel Ad Surya Pratama	0.002	0.000
Nama Responden	value=Nisami	0.002	0.000
Nama Responden	value=GHQZDHRZKZNYVH	0.002	0.000
Nama Responden	value=Karan Sembury	0.002	0.000
Nama Responden	value=A	0.005	0.000
Nama Responden	value=hans	0.002	0.000
Nama Responden	value=william	0.002	0.000
Nama Responden	value=Gilbert Lirimahua	0.002	0.000
Nama Responden	value=Cleora Hawae	0.002	0.000
Nama Responden	value=Claudia Labapapa	0.002	0.000
Nama Responden	value=Vitabella	0.002	0.000
Nama Responden	value=Magda Killy H	0.002	0.000
Nama Responden	value=Abel Pramudra Setiawan	0.000	0.059
Nama Responden	value=Idi	0.002	0.000

Gambar 9. Hasil Probabilitas Algoritma Naive Bayes

4. DISKUSI

Penelitian ini dibuat berdasarkan pengalaman peneliti sebelumnya yang sudah pernah melakukan penelitian yang menghasilkan beberapa sumber teori hasil atau analisis. Data-data dari penelitian sebelumnya tentu menjadi hal yang sangat dibutuhkan untuk melakukan penelitian ini. Data pendukung tersebut dapat digunakan penulis untuk melakukan pertimbangan dalam melakukan penelitian terhadap tingkat kepuasan konsumen Shopee menggunakan metode Naive Bayes dengan melakukan teknik data mining.

Pada penelitian terdahulu oleh Gustientiedina, Muhammad Siddik, dan Yenny Desnelita dengan judul "Penerapan Naive Bayes untuk Memprediksi Tingkat Kepuasan Mahasiswa Terhadap Pelayanan Akademis" menjelaskan dari hasil pengujian akhir yang dilakukan dari data sampel kuesioner dengan atribut indikator *tangible*, *reliability*, *responsiveness*, *assurance*, dan *emphaty* menggunakan metode naive bayes didapatkan tingkat akurasi sebesar 96,71% dengan nilai precision sebesar 96,15% dan nilai *recallnya* sebesar 98,43% [11].

Abdi, Sumijan, dan Gunadi dengan jurnalnya yang berjudul "Prediksi Tingkat Kepuasan dalam Pembelajaran Daring Menggunakan Algoritma Naive Bayes" menyimpulkan bahwa dari hasil pengujian akhir yang dilakukan dari data sampel kuesioner dengan menggunakan atribut komunikasi, suasana pembelajaran, penilaian mahasiswa, dan penyampaian materi menggunakan metode Naive Bayes yang didapatkan nilai akurasi sebesar 100% dengan nilai *precision* sebesar 100% dan nilai *recall* sebesar 100% [12].

Tri Herdiawan dan Castakan Agus dengan jurnal yang berjudul "Algoritma Naive Bayes untuk Prediksi Kepuasan Pelayanan Perekaman E-KTP" menjelaskan bahwa dari hasil pengujian yang dilakukan maka bisa ditarik kesimpulan dari seluruh pengujian dengan menggunakan metode Naive Bayes lebih akurat untuk data tingkat kepuasan pelayanan E-KTP di Kecamatan Batujajar dengan tingkat akurasi sebesar 91,70% dan nilai pengujian *f-measure* sebesar 93,92% [13].

Berdasarkan uraian beberapa penelitian sebelumnya, penulis tertarik untuk mengangkat topik yang sama dengan studi kasus yang berbeda dari penelitian sebelumnya. Dengan judul topik yang diangkat adalah "Teknik *Data Mining* menggunakan Algoritma Naive Bayes untuk Memprediksi Kepuasan Konsumen Shopee Indonesia di Kalangan Generasi Milenial". Diharapkan penelitian yang dibuat ini dapat meningkatkan akurasi penilaian terhadap pelayanan Shopee sehingga Shopee sendiri dapat melakukan evaluasi dan mempunyai standar dalam penilaian terhadap pelayanan.

5. KESIMPULAN DAN SARAN

Dari hasil penelitian dan analisis yang telah dilakukan oleh penulis terhadap klasifikasi data konsumen Shopee di Indonesia dalam melihat tingkat kepuasan konsumen, dapat ditarik kesimpulan bahwa konsumen merasa "Puas" saat menggunakan *platform* Shopee yang dapat dilihat dari berbagai aspek atribut. Hal ini juga dapat disimpulkan dari hasil klasifikasi menggunakan Algoritma Naive Bayes, dimana nilai probabilitas *class* "Tidak Puas" lebih kecil dari nilai probabilitas *class* "Puas". Setelah melakukan pengujian dengan menggunakan *tools* RapidMiner tingkat akurasi yang diperoleh sebesar 89.65% yang dimana termasuk dalam kategori "Puas". Jadi hasil akhir yang dapat disimpulkan adalah konsumen Shopee di Indonesia merasa puas berdasarkan pengalaman konsumen saat menggunakan *e-commerce* Shopee, yang dimana data ini juga dapat membantu pihak Perusahaan Shopee dalam menjaga dan meningkatkan kualitas untuk tetap menjaga kesetiaan konsumen Shopee. Adapun saran untuk pengembangan penelitian selanjutnya adalah sebagai berikut.

a. Adanya keterbatasan atribut pada kuesioner yang telah disebarakan sehingga atribut yang digunakan sedikit karena banyak terhapus pada

saat *cleaning data*, diharapkan untuk penelitian selanjutnya menggunakan atribut yang banyak.

b. Hanya terpatok pada satu *software* saja yaitu RapidMiner, akan lebih baik jika melakukan perbandingan terhadap beberapa *software* seperti *Orange*, *Weka*, *R* dan sebagainya demi mendapatkan hasil yang sangat akurat.

DAFTAR PUSTAKA

- [1] K. F. Sandy, "No TGenerasi Milenial Pengguna Transaksi *e-Commerce* Paling Banyak," <https://ekbis.sindonews.com/>, 2018. <https://ekbis.sindonews.com/berita/1357811/34/generasi-milenial-pengguna-transaksi-e-commerce-paling-banyak> (accessed Feb. 01, 2021).
- [2] Alwendi, "Penerapan E-Commerce Dalam Meningkatkan," *Manaj. Bisnis*, vol. 17, no. 3, pp. 317–325, 2020, [Online]. Available: <http://journal.undiknas.ac.id/index.php/magister-manajemen/>.
- [3] chusnul hajjah Murni, "PELANGGAN PADA ONLINE SHOP DI SHOPEE Pembimbing :," 2020.
- [4] A. Kurniawati and N. Ariyani, "Strategi Promosi Penjualan pada Marketplace Shopee," *Propaganda*, vol. 2, no. 1, pp. 65–79, 2022, doi: 10.37010/prop.v2i1.514.
- [5] Sambodo Rio Sasongko, "Faktor-Faktor Kepuasan Pelanggan Dan Loyalitas Pelanggan (Literature Review Manajemen Pemasaran)," *J. Ilmu Manaj. Terap.*, vol. 3, no. 1, pp. 104–114, 2021, doi: 10.31933/jimt.v3i1.707.
- [6] M. G. M. Wibowo, "Peran Ketidakpuasan Konsumen Dan Pencarian Variasi: Sebuah Studi Perpindahan Merek," *J. Ilmu Manaj.*, vol. 8, no. 2011, pp. 80–86, 2020.
- [7] N. I. Purnama, Y. Siswadi, and S. Mujiatun, "Model Perilaku Brand Switching Konsumen Dalam Pembelian Produk Otomotif," *J. Ilm. Manaj. Dan Bisnis*, vol. 22, no. 2, pp. 151–163, 2021, doi: 10.30596/jimb.v22i2.7305.
- [8] R. Rachman and R. N. Handayani, "Klasifikasi Algoritma Naive Bayes Dalam Memprediksi Tingkat Kelancaran Pembayaran Sewa Teras UMKM," *J. Inform.*, vol. 8, no. 2, pp. 111–122, 2021, doi: 10.31294/ji.v8i2.10494.
- [9] K. Sulastri, "Klasifikasi Naive Bayes pada Analisis Sentimen atas Penolakan Dibukanya Larangan Ekspor Benih Lobster," vol. 1, no. 2, pp. 68–75, 2020.
- [10] S. M. Sinaga, J. T. Hardinata, and M. Fauzan, "Implementasi Data Mining Clustering Tingkat Kepuasan Konsumen terhadap

- Pelayanan Go-Jek,” *Kesatria J. Penerapan Sist. Inf. (Komputer dan Manajemen)*, vol. 2, no. 2, pp. 118–124, 2021.
- [11] M. Siddik, Y. Desnelita, and Gustientiedina, “Penerapan Naïve Bayes untuk Memprediksi Tingkat Kepuasan Mahasiswa Terhadap Pelayanan Akademis,” *J. Infomedia*, vol. 2, no. 4, pp. 89–93, 2019, [Online]. Available: <http://e-jurnal.pnl.ac.id/index.php/infomedia/article/view/1892>.
- [12] A. R. Damanik, S. Sumijan, and G. W. Nurcahyo, “Prediksi Tingkat Kepuasan dalam Pembelajaran Daring Menggunakan Algoritma Naïve Bayes,” *J. Sistim Inf. dan Teknol.*, vol. 3, pp. 88–94, 2021, doi: 10.37034/jsisfotek.v3i3.137.
- [13] T. H. Apandi and C. A. Sugianto, “Algoritma Naive Bayes untuk Prediksi Kepuasan Pelayanan Perekaman e-KTP (Naive Bayes Algorithm for Satisfaction Prediction of e-ID,” *JUITA (Jurnal Inform. UMP)*, vol. 7, no. November, pp. 125–128, 2019.
- [14] A. A. Argasah and D. Gustian, “Data Mining Analysis To Determine Employee Salaries According To Needs Based On The K-Medoids Clustering Algorithm Analisis Data Mining Untuk Menentukan Gaji Karyawan Sesuai Penilaian Kemampuan Menggunakan Algoritma K-Medoids,” *Jutif*, vol. 3, no. 1, pp. 29–35, 2022.
- [15] A. Alwi, P. Studi, T. Informatika, F. Teknik, and U. M. Ponorogo, “the Concept of Naive Bayes and Its Simple Use for Prediction Final Konsep Naive Bayes Dan Penggunaannya Secara Sederhana,” vol. 3, no. 1, pp. 133–140, 2022.
- [16] Sofyan, “Implementation of Data Mining with Association Rule in Decision Making for Product Purchase Correlation Using the Apriori Algorithm,” *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2019.
- [17] F. S. Nugraha, M. J. Shidiq, and S. Rahayu, “Analisis Algoritma Klasifikasi Neural Network Untuk Diagnosis Penyakit Kanker Payudara,” *J. Pilar Nusa Mandiri*, vol. 15, no. 2, pp. 149–156, 2019, doi: 10.33480/pilar.v15i2.601.
- [18] P. R. Sihombing and A. M. Arsani, “Comparison of Machine Learning Methods in Classifying Poverty in Indonesia in 2018,” *J. Tek. Inform.*, vol. 2, no. 1, pp. 51–56, 2021, doi: 10.20884/1.jutif.2021.2.1.52.