

## Klasifikasi Kepuasan Pengguna Layanan Aplikasi Shopee Menggunakan Metode Decision Tree C4.5

Ballya Vicky Haekal<sup>1</sup>, lin Ernawati, S.Kom., M.Si.<sup>2</sup>, Nurul Chamidah, S.Kom, M.Kom.<sup>3</sup>  
<sup>1,2,3</sup>Program Studi Informatika, Universitas Pembangunan Nasional Veteran Jakarta  
<sup>1,2,3</sup>Jl. RS. Fatmawati Raya, Pd. Labu, Kec. Cilandak, Kota Depok, Jawa Barat 12450  
 email: <sup>1</sup>[ballyavh@upnvj.ac.id](mailto:ballyavh@upnvj.ac.id), <sup>2</sup>[iinernawati@upnvj.ac.id](mailto:iinernawati@upnvj.ac.id), <sup>3</sup>[nurul.chamidah@upnvj.ac.id](mailto:nurul.chamidah@upnvj.ac.id)

**Abstrak.** Seiring perkembangan teknologi, Saat ini sudah banyak sekali layanan aplikasi perdagangan *online* yang digunakan oleh masyarakat. Salah satu layanan aplikasi yang sering digunakan masyarakat adalah Shopee. Semakin banyak pengguna yang menggunakan aplikasi tersebut, tingkat kepuasan tiap pengguna pun semakin beragam. Dalam menganalisis tingkat kepuasan pengguna layanan aplikasi shopee, penelitian ini menggunakan data yang diambil dari kuesioner yang telah disebarluaskan dalam bentuk *google form*. Kuesioner tersebut disebarluaskan melalui beberapa media sosial seperti Twitter, Instagram, WhatsApp dan Shopee dengan total jumlah data yakni 184 data yang akhirnya menjadi 171 data setelah dilakukan *cleaning* data. Dari berbagai algoritma klasifikasi yang ada, penelitian ini menggunakan algoritma *Decision Tree C4.5* sebagai metode klasifikasinya. Dengan pembagian data sebesar 80% untuk data latih dan 20% untuk data uji. Hasil yang diperoleh pada penelitian ini terdapat nilai *Accuracy* sebesar 97%, nilai *Recall* sebesar 96.9%, nilai *Precision* sebesar 100% untuk *Class* “Ya” dan 66.6% untuk *Class* “Tidak”.

**Kata kunci:** Tingkat Kepuasan, Klasifikasi, *Decision Tree C4.5*, Shopee.

### 1. PENDAHULUAN

Seiring menyebarnya perkembangan era industri 4.0, teknologi informasi dan komunikasi terus berkembang mengikuti perubahan zaman. Tingkat mobilitas masyarakat dalam kegiatan sehari-hari menjadi semakin meningkat. Dengan tingkat mobilitas yang tinggi ini, masyarakat mulai membiasakan diri dengan gaya hidup yang serba instan dalam memenuhi kebutuhan hidup. Hal ini membuat berbagai penyedia layanan bermunculan untuk membantu masyarakat.

Saat ini di Indonesia, telah ada berbagai macam penyedia layanan e-commerce. Salah satu layanan e-commerce yang ada di Indonesia adalah Shopee. Shopee adalah salah satu situs dari sekian banyak *E-Commerce* pertama diluncurkan pada tahun 2015, dibawah kepemimpinan Chris Feng. Berkantor pusat di Singapura, dibawah naungan SEA Group yang telah berdiri sejak tahun 2009 dibawah pimpinan Forrest li.[1] Banyaknya promo menjadi salah satu alasan Shopee banyak diminati masyarakat. Shopee menyediakan berbagai macam promo, mulai dari diskon produk, *voucher cashback*, *voucher* gratis ongkos kirim dan masih banyak lagi. Hal ini tentunya menarik minat pembeli untuk melakukan transaksi dengan menggunakan aplikasi Shopee. [2]

Layanan merupakan usaha dalam memenuhi kebutuhan internal dan eksternal untuk pelanggan sesuai prosedur dengan konsisten. Dalam pelayanan, penyedia jasa diharuskan untuk memiliki kemampuan untuk mengerti apa yang diinginkan pelanggan. [3] Layanan yang baik dapat dilihat dari tingkat kepuasan pelanggan. Tingkat kepuasan pengguna yaitu dimana anggapan terhadap layanan sesuai dengan sesuatu yang diinginkan oleh pengguna. Prakiraan dan keyakinan konsumen menjadi suatu harapan konsumen dengan sesuatu yang diterimanya setelah mencoba suatu layanan. Setiap pengguna memiliki berbagai tingkat kepuasan, ketika layanan tersebut sesuai dengan apa yang diinginkannya maka pengguna tersebut akan merasa terpuaskan dengan hasil tersebut. Sebaliknya, jika layanan tersebut tidak sesuai dengan apa yang diinginkannya maka pengguna tersebut akan merasa tidak puas dengan hasil tersebut. [4]

Dengan banyaknya pengguna aplikasi Shopee, tentunya tidak semua pengguna merasa puas dengan layanan aplikasi tersebut. Terlebih lagi, Shopee bukan satu-satunya penyedia layanan *e-commerce* yang ada. Hal ini membuat Shopee membutuhkan analisis terhadap kepuasan penggunanya. Oleh karena itu, data mining diperlukan untuk menganalisis kepuasan pengguna Shopee. [5] Data mining merupakan metode analisis dari

proses pencarian pengetahuan pada basis data atau biasa disebut *knowledge discovery in databases* (KDD). Pengetahuan tentang masalah dapat berbentuk model data, atau hubungan antara data yang valid atau yang sebelumnya tidak diketahui. Data mining adalah proses menemukan pengetahuan, pola, dan hubungan antar suatu data secara otomatis. Pengetahuan dapat ditemukan dengan lima proses berurutan yaitu seleksi, pra pemrosesan, transformasi, data mining, dan interpretasi atau evaluasi.[6]

Salah satu metode data mining adalah klasifikasi. Klasifikasi merupakan metode yang digunakan untuk menetapkan *record* data baru ke salah satu kelas atau kategori yang telah ditentukan sebelumnya. Dengan kata lain klasifikasi juga berguna untuk memprediksi label di tiap kategori kelas.[7] Salah satu algoritma klasifikasi adalah *Decision Tree*. *Decision tree* merupakan metode klasifikasi data yang mudah dimengerti. *Decision tree* membantu dalam penggambaran aturan yang mudah dipahami dan juga dapat dikembangkan menjadi suatu basis data. *Decision tree* dapat digunakan untuk melakukan eksplorasi data, mencari hubungan yang tidak diketahui atau tersembunyi diantara variabel yang masuk dan target. [8]

Terdapat banyak penelitian terdahulu yang bersangkutan dengan dilakukannya penelitian ini seperti penelitian [7] yang menerapkan metode klasifikasi pada 300 data restoran cepat saji di kota bogor dengan kategori: Sangat puas, puas, tidak puas dan sangat tidak puas. Penelitian ini memperoleh hasil berupa model dengan tingkat akurasi sebesar 91% dengan nilai presisi 92.21% pada prediksi puas dan 90.91% pada prediksi tidak puas.

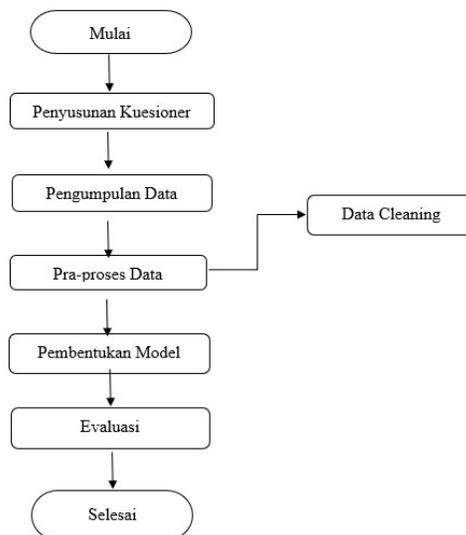
Pada penelitian [4] dilakukan penelitian menggunakan algoritma *Decision Tree* C4.5 untuk menentukan faktor-faktor yang mempengaruhi kepuasan pelanggan restoran. Hasil dari penelitian ini yaitu pelanggan akan merasa puas apabila pelayanan dirasa lumayan, produk dirasa lumayan, fasilitas dirasa baik dan pelayan dirasa ramah.

Pada penelitian [8] dilakukan penerapan algoritma C4.5 untuk memprediksi siswa SMA Bina Bangsa Mandiri Bogor yang akan mendapatkan dana bantuan pendidikan. Hasil yang diperoleh penelitian ini berupa model dengan tingkat akurasi sebesar 98.80% dengan presisi sebesar 98.02%.

Pada penelitian [9] dilakukan klasifikasi dengan menggunakan *algoritma Naive Bayes* terhadap sentimen aplikasi Shopee. Data yang digunakan dari penelitian ini diambil dari review aplikasi shopee yang ada pada *PlayStore*. Penelitian ini menghasilkan akurasi sebesar 71.50% dengan nilai Area Under Curve (AUC) sebesar 0.500.

Dalam penelitian ini, metode data mining yang digunakan adalah klasifikasi. Klasifikasi merupakan pengelompokan berdasarkan hubungan antara variabel kriteria dengan variabel target. Dalam proses klasifikasi terdapat beberapa algoritma yang dapat digunakan. Pada penelitian ini, algoritma yang akan digunakan yaitu algoritma *Decision Tree* C4.5. Pada penelitian ini akan dilakukan penerapan salah satu metode data mining yaitu metode klasifikasi untuk mendapatkan hasil klasifikasi untuk mengetahui tingkat kebugaran jasmani dari hasil pengukuran dengan menggunakan algoritma *Decision Tree* C4.5 dengan hasil berupa hasil klasifikasi.

## 2. METODE PENELITIAN



Gambar 1. Metode Penelitian

### 2.1 Penyusunan Kuisisioner

Pada langkah ini dilakukan penyusunan kuisisioner. Kuisisioner disusun berdasarkan dengan sumber referensi-referensi yang sudah didapat sebelumnya.

Tabel 1. Data Hasil Kuesioner

Variabel	1	2	3	4	....	184
Nama	Dhe...	Muh...	Sul...	Jef...	....	Ani...
Email	Dhe...@gmail.com	hai...@gmail.com	sul...@gmail.com	Jef...@gmail.com	....	ani...@gmail.com
Umur	18 - 20 Tahun	18 - 20 Tahun	21 - 30 Tahun	21 - 30 Tahun	....	21 - 30 Tahun
Jenis Kelamin	Perempuan	Laki - laki	Laki - laki	Laki - laki	....	Perempuan
X1	5	5	5	4	....	3
X2	5	5	5	4	....	3
X3	4	5	5	5	....	3
X4	3	5	5	4	....	3
X5	2	5	5	5	....	3
X6	1	5	5	5	....	3
X7	4	3	4	4	....	3
X8	4	4	4	5	....	3
X9	4	3	4	5	....	3
X10	2	3	4	5	....	3
X11	3	4	3	5	....	3
X12	4	3	3	4	....	3
X13	5	3	3	4	....	3
X14	5	4	4	4	....	3
X15	4	4	4	4	....	3
X16	2	4	4	4	....	3
X17	4	5	3	4	....	3
X18	3	5	3	5	....	3
X19	3	4	3	3	....	3
X20	3	3	3	3	....	3
X21	2	3	3	4	....	3
Class	Ya	Tidak	Ya	Tidak	....	Ya

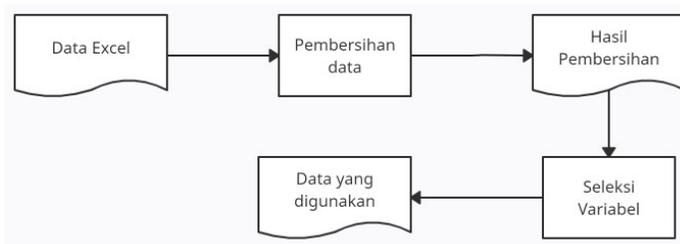
### 2.2 Pengumpulan Data

Pada tahap ini dilakukan pengumpulan data untuk penelitian. Data yang digunakan berasal dari kuesioner dalam bentuk *google form*. Data yang telah dikumpulkan dari kuesioner tersebut berjumlah 184 data dengan 171 data berada pada *Class* “Ya” dan 13 data berada pada *Class* “Tidak”. Data yang digunakan pada penelitian ini didapatkan setelah pengumpulan data melalui kuesioner yang telah disebar. Kuesioner tersebut terdiri dari data pribadi responden beserta 22 pertanyaan yang dikelompokkan menjadi 8 kategori yaitu data pribadi, reabilitas, *information quality*, *interaction quality*, *efficiency*, *contact*, *user satisfication*, *web design* and *layout*. Masing-masing pertanyaan menggunakan skala likert 1 sampai 5 sebagai pilihan jawabannya.

*Class* yang akan digunakan pada penelitian ini diambil langsung melalui pertanyaan terakhir yang ada pada pertanyaan yang dibuat, dengan jumlah *class* sebanyak 2 *class* yakni “Ya” dan “Tidak”. Hasil pengumpulan data dengan kuesioner dapat dilihat pada Tabel 1.

### 2.3 Pra-Proses Data

Tahap pra-proses data untuk menjadikan data menjadi suatu data yang dapat digunakan dan mudah untuk dibaca (*readable*). Tahap-tahap yang ada pada pra-proses penelitian ini sebagai berikut.



Gambar 2. Pra-Proses Data

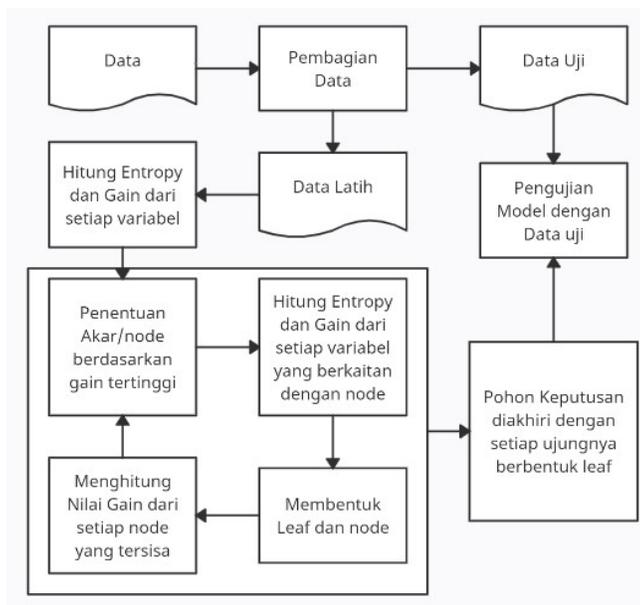
Berdasarkan Gambar 2 di atas, data yang telah dikumpulkan kemudian dilakukan pra-proses dengan membersihkan data dari data *duplicate* dan juga melakukan seleksi variabel untuk menghilangkan variabel yang tidak terpakai. Pada penelitian ini pra-proses data yang dilakukan adalah *cleaning data* dan *drop duplicate data*. Setelah dilakukan *cleaning data*, Langkah selanjutnya yaitu melakukan *drop duplicate data*. *Drop duplicate data* disini digunakan untuk menghapus data yang sama agar tidak terjadi data yang berulang.

Tabel 2. Data Setelah Pra-proses

Variabel	1	2	3	4	....	171
X1	5	5	5	4	....	3
X2	5	5	5	4	....	3
X3	4	5	5	5	....	3
X4	3	5	5	4	....	3
X5	2	5	5	5	....	3
X6	1	5	5	5	....	3
X7	4	3	4	4	....	3
X8	4	4	4	5	....	3
X9	4	3	4	5	....	3
X10	2	3	4	5	....	3
X11	3	4	3	5	....	3
X12	4	3	3	4	....	3
X13	5	3	3	4	....	3
X14	5	4	4	4	....	3
X15	4	4	4	4	....	3
X16	2	4	4	4	....	3
X17	4	5	3	4	....	3
X18	3	5	3	5	....	3
X19	3	4	3	3	....	3
X20	3	3	3	3	....	3
X21	2	3	3	4	....	3
Class	Ya	Tidak	Ya	Tidak	....	Ya

Tahapan berikutnya adalah melakukan Seleksi Variabel untuk menghilangkan data yang tidak diperlukan yaitu nama, email, umur dan jenis kelamin. Untuk variabel umur, pada penelitian ini variabel tersebut tidak digunakan karena  $\pm 90\%$  dari data yang ada berada pada umur 18-30 tahun. Setelah dilakukan pra-proses data dapat dilihat pada Tabel 2.

## 2.4 Pembentukan Model



Gambar 3. Pembentukan Model

*Decision tree* merupakan metode klasifikasi data yang mudah dimengerti. *Decision tree* membantu dalam penggambaran aturan yang mudah dipahami dan juga dapat dikembangkan menjadi suatu basis data. *Decision tree* dapat digunakan untuk melakukan eksplorasi data, mencari hubungan yang tidak diketahui atau tersembunyi diantara variabel yang masuk dan target. [9]

Berdasarkan Gambar 3 di atas, pembentukan model dilakukan dengan algoritma *Decision Tree C4.5*. Dengan menggunakan pembagian data sebesar 80% digunakan sebagai data latih dan 20% untuk data uji.

Setelah itu hitung nilai *Entropy* total beserta nilai *Entropy* seluruh atributnya menggunakan rumus *Entropy*. *Entropy* merupakan nilai ketidakmurnian (*impurity*) suatu atribut. Semakin kecil nilai *entropy* maka akan semakin baik digunakan untuk mengekstrak suatu kelas. Adapun rumus untuk mencari *entropy*. [10]

$$Entropy(S) = - \sum_{i=1}^n p_i * \log_2 p_i$$

Dimana:

S: himpunan kasus

n: jumlah partisi S

Pi: proporsi dari  $S_i$  terhadap S

Nilai *entropy* ini yang nantinya akan digunakan untuk menghitung nilai *gain*. Nilai *gain* digunakan untuk memisahkan objek. Adapun rumus untuk mencari *gain*. [11]

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i)$$

Dimana:

S: himpunan kasus

A: variabel

n: jumlah partisi variabel A

|S<sub>i</sub> |: jumlah kasus pada partisi ke-1

|S|: jumlah kasus dalam S

Atribut dengan nilai *gain* terbesar kemudian menjadi akar (node) dilanjutkan dengan penghitungan kembali setiap atribut lainnya dengan cara yang sama untuk menentukan cabang-cabang berikutnya.

### 2.5 Evaluasi

Pada tahap ini, dilakukan perhitungan tingkat akurasi data dengan menggunakan *confusion matrix*, dari *Confusion Matrix* tersebut akan dihitung nilai *Accuracy*, *Recall* dan *Precision* dengan rumus sebagai berikut.

$$accuracy = \frac{TP+TN}{TP+FP+TN+FN} \dots\dots(3)$$

$$Recall = \frac{TP}{TP+FN} \dots\dots(4)$$

$$Precision = \frac{TP}{TP+FP} \dots\dots(5)$$

## 3. HASIL DAN PEMBAHASAN

### 3.1 Pembentukan Model

Pada tahap ini data akan dibagi menjadi data latih dan data uji. Metode pembagian data yang digunakan pada penelitian ini adalah hold out validation. Hold out validation adalah pembagian data secara acak dengan komposisi 80% untuk data latih dan 20% sisanya untuk data uji. Pembagian data ini juga dilakukan dengan menyeimbangkan jumlah data pada tiap class agar terbagi secara adil untuk data latih dan data uji. Dengan persentase pembagian data tersebut, 171 data tersebut dibagi menjadi 136 data sebagai data latih meliputi 127 data berada pada *Class* “Ya” dan 9 data berada pada *Class* “Tidak” dan 35 data sebagai data uji meliputi 33 data berada pada *Class* “Ya” dan 2 data berada pada *Class* “Tidak”.

Tahapan pertama yang dilakukan adalah menghitung nilai *Entropy* dari semua kriteria yang ada. Lalu hitung nilai *Gain* dari setiap atribut. Langkah ini dilakukan pada semua variabel yang ada. Setelah nilai *Gain* dari setiap variabel ditemukan, selanjutnya adalah menentukan akar berdasarkan nilai *Gain* tertinggi.

Langkah kedua adalah mencari *node* kanan dan kiri pohon keputusan yang dilakukan dengan menganalisis atribut-atribut dari variabel tersebut. Langkah ketiga adalah menghitung Kembali nilai *Entropy* dan nilai *Gain* dari setiap variabel berdasarkan data yang ada.

Setelah nilai *Gain* dari setiap variabel ditemukan, Langkah keempat adalah menentukan node dengan variabel yang memiliki nilai *Gain* tertinggi. Langkah kelima ialah mengulangi Langkah pertama hingga Langkah keempat sampai tidak ada lagi variabel yang dapat digunakan sebagai cabang.

### 3.2 Evaluasi

Setelah model selesai dibangun, Langkah selanjutnya adalah menguji model tersebut dengan menggunakan 35 data uji yang meliputi 33 data berada pada *Class* “Ya” dan 2 data berada pada *Class* “Tidak”. Hasil dari pengujian tersebut dapat dilihat pada Tabel 3 dibawah ini.

**Tabel 3.** Confusion Matrix

Class		Prediksi	
		Ya	Tidak
Aktual	Ya	32	1
	Tidak	0	2

Dengan nilai pada Tabel tersebut, dapat dilakukan perhitungan *Accuracy* dengan persamaan (3), *Recall* dengan persamaan (4), dan *Precision* dengan persamaan (5) sebagai berikut:

$$Accuracy = \frac{32 + 2}{32 + 0 + 2 + 1} = \frac{34}{35} = 0.97$$

$$Recall = \frac{32 + 1}{32 + 1} = \frac{33}{33} = 0.969$$

$$Precision_{Ya} = \frac{32}{32 + 0} = \frac{32}{32} = 1$$

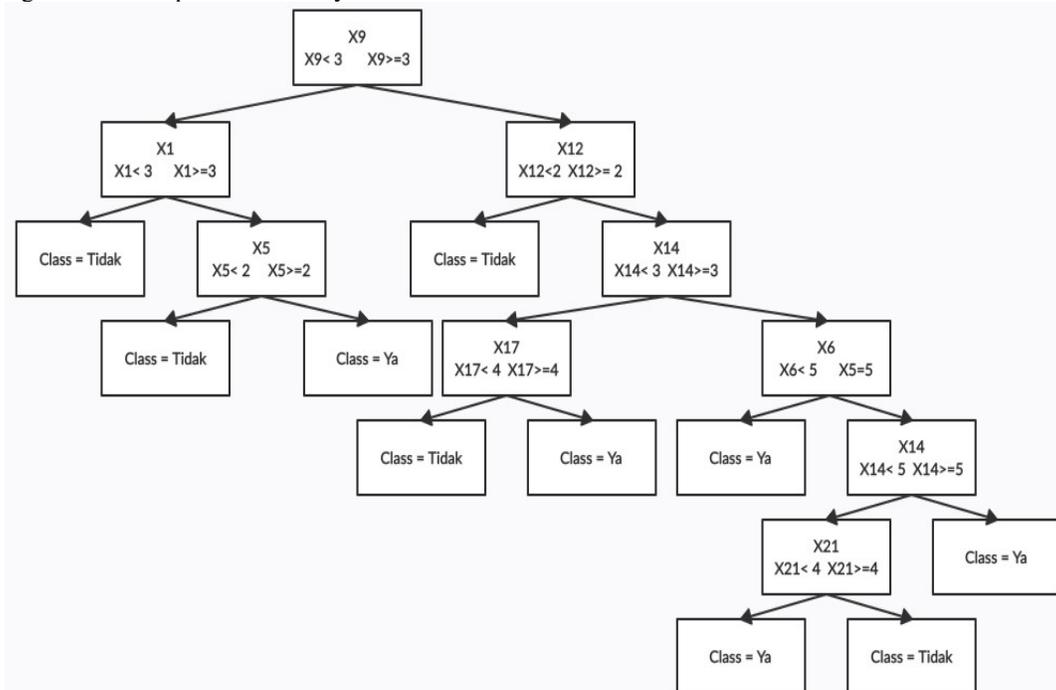
$$Precision_{Tidak} = \frac{2}{2 + 1} = \frac{2}{3} = 0.666$$

Hasil dari perhitungan *Confusion Matrix* dari seluruh percobaan pengujian dapat dilihat pada Tabel 4 dibawah ini.

**Tabel 4.** Hasil *Confusion Matrix* Seluruh Percobaan

No	Accuracy	Recall	Precision (Ya)	Precision (Tidak)
1	0.8857	0.9090	0.9677	0.25
2	0.9142	0.9393	0.9687	0.3333
3	0.8857	0.9090	0.9677	0.25
4	0.9142	0.9090	1.0	0.4
5	0.9428	0.9696	0.9696	0.5
6	0.9428	0.9696	0.9696	0.5
7	0.9142	0.9393	0.9687	0.3333
8	0.9714	0.9696	1.0	0.6666
9	0.8857	0.9090	0.9677	0.25

Berdasarkan Tabel 4 diatas, dapat dilihat percobaan yang memiliki tingkat *Accuracy* yang paling tinggi adalah percobaan kedelapan. Sehingga, model yang digunakan pada percobaan kedelapan merupakan model yang paling baik diantara percobaan lainnya.



**Gambar 4.** Hasil Pohon Keputusan

### 3.3 Hasil Akhir

Berdasarkan Tabel 4 dapat dilihat percobaan yang memiliki tingkat *Accuracy* yang paling tinggi adalah percobaan kedelapan. Sehingga, model yang digunakan pada percobaan kedelapan merupakan model yang paling baik diantara percobaan lainnya. Berikut ini merupakan hasil pohon keputusan terbaik yang telah dibangun, dapat dilihat pada Gambar 4.

Berdasarkan model yang telah dibangun dengan melihat pohon keputusan pada Gambar 4, menunjukkan bahwa tidak semua variabel berpengaruh dalam pembangunan model yang telah dibuat. Variabel-variabel yang berpengaruh dalam pembangunan model yang sudah dibangun dapat dilihat pada Tabel 5 dibawah ini.

**Tabel 5.** Variabel yang Mempengaruhi Hasil Klasifikasi

Variabel Pengganti	Variabel
X9	Harga di Shopee lebih relatif murah dibanding dengan beli di toko offline.
X1	Anda memilih shopee karena sering mengadakan promo setiap bulannya.
X12	Shopee mudah untuk digunakan.
X5	Program Rp10.000 membuat saya tertarik untuk mengikuti event tersebut, sehingga saya memilih Shopee.
X14	Pemilihan jasa pengiriman lebih beragam
X17	Dengan adanya fitur Shopeepay memudahkan proses pembayaran.
X6	Shopee sering mengadakan giveaway yang membuat saya tertarik menggunakan Shopee.
X21	Customer service sangat membantu ketika mengalami kendala.

Dari Gambar 4 Hasil Pohon Keputusan diperoleh sebanyak 10 aturan klasifikasi, berikut aturan klasifikasi yang terbentuk berdasarkan model yang telah dibuat.

1. Jika  $X9 < 3$  dan  $X1 < 3$  maka akan berada pada *Class* “Tidak”.
2. Jika  $X9 \geq 3$  dan  $X12 < 2$  maka akan berada pada *Class* “Tidak”.
3. Jika  $X9 < 3$ ,  $X1 \geq 3$  dan  $X5 < 2$  maka akan berada pada *Class* “Tidak”.
4. Jika  $X9 < 3$ ,  $X1 \geq 3$  dan  $X5 \geq 2$  maka akan berada pada *Class* “Ya”.
5. Jika  $X9 \geq 3$ ,  $X12 \geq 2$ ,  $X14 < 3$  dan  $X17 < 4$  maka akan berada pada *Class* “Tidak”.
6. Jika  $X9 \geq 3$ ,  $X12 \geq 2$ ,  $X14 < 3$  dan  $X17 \geq 4$  maka akan berada pada *Class* “Ya”.
7. Jika  $X9 \geq 3$ ,  $X12 \geq 2$ ,  $X14 \geq 3$  dan  $X6 < 5$  maka akan berada pada *Class* “Ya”.
8. Jika  $X9 \geq 3$ ,  $X12 \geq 2$ ,  $X14 \geq 3$ ,  $X6 \geq 5$  dan  $X14 = 5$  maka akan berada pada *Class* “Ya”.
9. Jika  $X9 \geq 3$ ,  $X12 \geq 2$ ,  $X14 \geq 3$ ,  $X6 \geq 5$ ,  $X14 < 5$  dan  $X21 < 4$  maka akan berada pada *Class* “Tidak”.
10. Jika  $X9 \geq 3$ ,  $X12 \geq 2$ ,  $X14 \geq 3$ ,  $X6 \geq 5$ ,  $X14 < 5$  dan  $X21 \geq 4$  maka akan berada pada *Class* “Ya”.

Model diatas menghasilkan performa berupa nilai *Accuracy* sebesar 0.9714, nilai *Recall* sebesar 0.9696, nilai *Precision* “Ya” sebesar 1 dan nilai *Precision* “Tidak” sebesar 0.6666.

## 4. KESIMPULAN DAN SARAN

Berdasarkan penelitian yang telah dilakukan dapat disimpulkan bahwa dengan menyebarkan kuesioner untuk pengguna aplikasi shopee untuk mengetahui kepuasan pengguna layanan aplikasi shopee dengan variabel berjumlah 21 variabel dan 1 variabel *Class*, yang dipra-proses dengan Seleksi Variabel dan *Data Cleaning* dengan pembagian data 80% untuk membangun model dan 20% untuk menguji model. Penggunaan metode klasifikasi Decision Tree C4.5 menghasilkan nilai *Accuracy* sebesar 0.97, nilai *Recall* sebesar 0.969, nilai *Precision* sebesar 1 untuk *Class* “Ya” dan 0.666 untuk *Class* “Tidak”. Berdasarkan hasil tersebut, performa dari algoritma Decision Tree C4.5 sudah cukup baik dalam mengklasifikasi tingkat kepuasan pengguna shopee. Variabel yang paling berpengaruh terhadap pembangunan model kepuasan pengguna layanan aplikasi shopee adalah “Harga di Shopee lebih relatif murah dibanding dengan e-commerce lain.”.

Saran untuk penelitian ini diantaranya menyebarkan kuesioner lebih banyak agar data yang digunakan lebih bervariasi. Hal ini juga dapat membantu data agar jumlah data dari setiap class lebih seimbang. Sehingga nilai *Precision* yang dihasilkan lebih besar dibanding penelitian ini. Serta diharapkan dapat mengubah variabel pertanyaan sehingga variabel yang mempengaruhi hasil klasifikasi dapat lebih bervariasi.

## 5. DAFTAR PUSTAKA

- [1] Techinasia (22 November 2018). *Shopee sales surge helps Sea hit record revenue, but net loss widens*, dari <https://www.techinasia.com/shopee-sales-surge-helps-sea-hit-record-revenue>
- [2] Widyanita, Fika Ayu. (2018). “Analisis Pengaruh Kualitas Pelayanan E-Commerce Shopee Terhadap Kepuasan Konsumen Shopee Indonesia Pada Mahasiswa FE UII Pengguna Shopee”. Skripsi. Fakultas Ekonomi, Manajemen, Universitas Islam Indonesia, Yogyakarta.
- [3] Arista Atmadjati. (2018). *Layanan Prima dalam Praktik Saat Ini*. Deepublish.
- [4] Shiddiq, Ahmad, et. al. (2018). ‘Analisa Kepuasan Konsumen Menggunakan Klasifikasi Decision Tree di Restoran Dapur Solo (Cabang Kediri)’, jurnal: *Generation Journal / Vol.2 No.1 Januari 2018 / e-ISSN: 2549-2233 / p-ISSN: 2580-4952*
- [5] Suyanto. (2017). *Data Mining Untuk Klasifikasi dan Klasterisasi Data*. Penerbit Informatika.
- [6] Prasetyowati, E. (2017). *DATA MINING Pengelompokan Data untuk Informasi dan Evaluasi*. Duta Media Publishing.
- [7] Dhika, Harry, et. al. (2016) ‘Implementasi Algoritma C4.5 Terhadap Kepuasan Pelanggan’, jurnal: *Prosiding SNaPP2016 Sains dan Teknologi Vol 6, No.1, Th, 2016 ISSN 2089-3582, EISSN 2303-2480*.
- [8] Hendrian, Senna. (2018). ‘Algoritma Klasifikasi Data Mining untuk Memprediksi Siswa dalam Memperoleh Bantuan Dana Pendidikan’, jurnal : *Faktor Exacta 11 (3): 266-274, 2018, p-ISSN: 1979-276X, e- ISSN: 2502-339X*
- [9] Masripah, Siti., Lila Dini. (2020). ‘Algoritma Klasifikasi Naïve Bayes untuk Analisa Sentimen Aplikasi Shopee’, jurnal: *SWABUMI, Vol.8 No.2 September 2020, pp. 114~117, ISSN: 2355-990X, E-ISSN: 2549-5178*
- [10] Oktafianto. (2016). ‘Analisis Kepuasan Mahasiswa Terhadap Pelayanan Akademik Menggunakan Metode Algoritma C4.5 (Studi Kasus: Stmik Pringsewu)’, jurnal: *TIM Darmajaya Vol. 02 No. 01 Mei 2016, ISSN: 2442-5567, E-ISSN: 2443-289X*
- [11] Nofriansyah, Dicky. 2017. *Algoritma Data Mining dan Pengujiannya*. STMIK Triguna Medan.