

**IMPLEMENTASI DATA MINING MENGGUNAKAN ALGORITMA C4.5  
PADA KLASIFIKASI PENJUALAN HIJAB**

**SKRIPSI**

**OLEH  
FARIDATUL HUSNA  
NIM. 17610069**



**PROGRAM STUDI MATEMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM  
MALANG  
2021**

**IMPLEMENTASI DATA MINING MENGGUNAKAN ALGORITMA C4.5  
PADA KLASIFIKASI PENJUALAN HIJAB**

**SKRIPSI**

**Diajukan Kepada  
Fakultas Sains dan teknologi  
Universitas Islam Negeri Maulana Malik Ibrahim Malang  
untuk Memenuhi Salah Satu Persyaratan dalam  
Memperoleh Gelar Sarjana Matematika (S.Mat)**

**Oleh  
Faridatul Husna  
NIM. 17610069**

**PROGRAM STUDI MATEMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM  
MALANG  
2021**

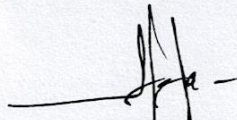
**IMPLEMENTASI DATA MINING MENGGUNAKAN ALGORITMA C4.5  
PADA KLASIFIKASI PENJUALAN HIJAB**

**SKRIPSI**

**Oleh  
Faridatul Husna  
NIM. 17610069**

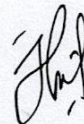
Telah Diperiksa dan Disetujui untuk Diuji  
Tanggal 10 Desember 2021

Pembimbing I,



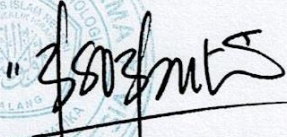
Dr. Hairul Rahman, M.Si  
NIP. 19800429 200604 1 003

Pembimbing II,



Juhari, M.Si  
NIDT. 19840209 20160801 1 055

Mengetahui,  
Ketua Program Studi Matematika

  
Dr. Elly Susanti, M.Sc  
NIP. 19741129 200012 2 005

**IMPLEMENTASI DATA MINING MENGGUNAKAN ALGORITMA C4.5  
PADA KLASIFIKASI PENJUALAN HIJAB**

**SKRIPSI**

**Oleh**  
**Faridatul Husna**  
**NIM. 17610069**

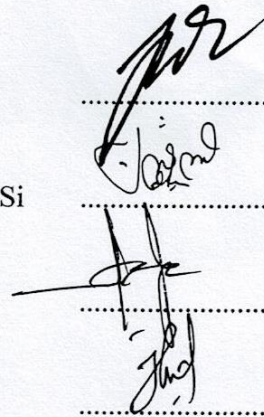
Telah Dipertahankan di Depan Dewan Penguji Skripsi  
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan  
Untuk Memperoleh Gelar Sarjana Matematika (S.Mat)  
Tanggal 28 Desember 2021

Penguji Utama : Muhammad Khudzaifah, M.Si

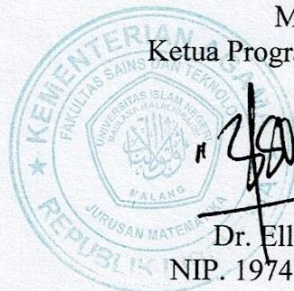
Ketua Penguji : Ria Dhea Layla Nur Karisma, M.Si

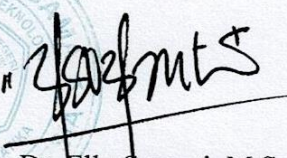
Sekretaris Penguji : Dr. Hairur Rahman, M.Si

Anggota Penguji : Juhari, M.Si



Mengetahui,  
Ketua Program Studi Matematika



  
Dr. Elly Susanti, M.Sc  
NIP. 19741129 200012 2 005

## PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Faridatul Husna  
NIM : 17610069  
Program Studi : Matematika  
Fakultas : Sains dan Teknologi  
Judul Skripsi : Implementasi Data Mining Menggunakan Algoritma C4.5 pada  
Klasifikasi Penjualan Hijab

menyatakan dengan sebenarnya bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya sendiri, bukan merupakan pengambilan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar rujukan. Apabila di kemudian hari terbukti atau dapat dibuktikan skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 10 Desember 2021

Yang membuat pernyataan,



Faridatul Husna

NIM. 17610069

## **MOTO**

“Barangsiapa tidak mensyukuri yang sedikit, maka ia tidak akan mampu mensyukuri yang banyak.”

(HR. Ahmad)

## **PERSEMBAHAN**

Dengan rasa syukur kepada Allah Swt, skripsi dipersembahkan kepada:  
Ayahanda Suyanto (Alm) dan Ibunda Siti Asiyah tercinta, atas ketulusan doanya, mujahadahnya, cintanya, kasih sayangnya, perhatian, dan pengorbanannya. Serta adik tersayang Muhammad Rofiqul Minan yang telah memberikan dukungan dan senantiasa memberikan motivasi untuk terus melangkah dan menebar kebaikan di bumi ini.

## KATA PENGANTAR

*Assalamu 'alaikum Warahmatullahi Wabarakatuh*

Segala puji bagi Allah Subhanahu Wa Ta'ala yang telah melimpahkan rahmat, taufik, serta hidayah-Nya, sehingga penulis masih diberi kesehatan dan kesempatan untuk dapat menyelesaikan skripsi yang berjudul “Implementasi Data Mining Menggunakan Algoritma C4.5 Pada Klasifikasi Penjualan Hijab” sebagai salah satu syarat untuk memperoleh gelar sarjana dalam bidang Matematika di Fakultas Sains dan teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Shalawat dan salamt semoga senantiasa tercurahkan kepada junjungan kita Nabi Muhammad Shallallahu ‘Alaihi Wasallam yang telah membawa kita dari zaman jahiliah menuju zaman islamiah.

Dalam penyusunann skripsi ini penulis menerima bimbingan dan bantuan dari berbagai pihak, oleh karena itu penulis mengucapkan terima kasih kepada :

1. Dr. Hairur Rahman, M.Si selaku Dosen Pembimbing I yang telah memberikan banyak ilmu, arahan, masukan, dan nasihat kepada penulis.
2. Juhari, M.Si selaku Dosen Pembimbing II yang telah memberikan banyak ilmu, arahan, masukan, dan nasihat kepada penulis.
3. Muhammad Khudzaifah, M.Si selaku Penguji Utama yang telah memberikan banyak ilmu, arahan, masukan, dan nasihat kepada penulis.
4. Ria Dhea Layla Nur Karisma, M.Si selaku Ketua Penguji yang telah memberikan banyak ilmu, arahan, masukan, dan nasihat kepada penulis.
5. Hisyam Fahmi, M.Kom selaku Dosen Konsentrasi Komputasi yang telah memberikan banyak ilmu, arahan, masukan, dan nasihat kepada penulis.



6. Seluruh sivitas akademika Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang.
7. Ayahanda (Alm) dan Ibunda yang selalu memberikan doa terbaik kepada penulis.
8. Seluruh teman-teman di Program Studi Matematika angkatan 2017.

Semoga Allah Swt melimpahkan rahmat dan karunia-Nya kepada kita semua. Penulis berharap skripsi ini dapat bermanfaat tidak hanya bagi penulis, tetapi juga bermanfaat bagi semua pihak yang membacanya. *Aamiin Allahumma Aamiin.*

*Wassalamu'alaikum Warahmatullahi Wabarakatuh.*

Malang, 30 November 2021

Penulis

## DAFTAR ISI

<b>HALAMAN JUDUL</b>	
<b>HALAMAN PENGAJUAN</b>	
<b>HALAMAN PESETUJUAN</b>	
<b>HALAMAN PENGESAHAN</b>	
<b>HALAMAN PERNYATAAN KEASLIAN TULISAN</b>	
<b>HALAMAN MOTO</b>	
<b>HALAMAN PERSEMBAHAN</b>	
<b>KATA PENGANTAR</b> .....	viii
<b>DAFTAR ISI</b> .....	x
<b>DAFTAR TABEL</b> .....	xii
<b>DAFTAR GAMBAR</b> .....	xiii
<b>ABSTRAK</b> .....	xiv
<b>ABSTRACT</b> .....	xv
<b>مخلص</b> .....	xvi
<b>BAB I PENDAHULUAN</b>	
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah .....	3
1.3 Tujuan Penelitian.....	4
1.4 Manfaat Penelitian.....	4
1.5 Batasan Masalah.....	4
1.6 Sistematika Penulisan.....	5
<b>BAB II KAJIAN PUSTAKA</b>	
2.1 Decision Tree.....	6
2.2 Klasifikasi.....	7
2.3 Algoritma C4.5 .....	7
2.4 <i>Confusion Matrix</i> .....	14
2.5 <i>K-Fold Cross Validation</i> .....	16
2.6 Kajian Keislaman .....	16
<b>BAB III METODE PENELITIAN</b>	
3.1 Sumber Data .....	19
3.2 Variabel Penelitian .....	19
3.3 Langkah-Langkah Penelitian.....	20

## **BAB IV PEMBAHASAN**

4.1 Analisis Data .....	22
4.1.1 Pembersihan Data .....	22
4.1.2 Transformasi Data .....	23
4.1.3 Proses Pembentukan <i>Decision Tree</i> Menggunakan Algoritma C4.5 ..	26
4.2 Uji Akurasi Algoritma C4.5 .....	38
4.3 <i>K-Fold Cross Validation</i> .....	40
4.4 Konsep Decision Tree untuk Mengatur Proses Jual Beli dalam Islam .....	40

## **BAB V PENUTUP**

5.1 Kesimpulan.....	43
5.2 Saran.....	43

## **DAFTAR PUSTAKA**

## **LAMPIRAN**

## **RIWAYAT HIDUP**

## DAFTAR TABEL

Tabel 2.1 Data Cuaca dan Keputusan untuk Main atau Tidak.....	8
Tabel 2.2 <i>Confusion Matrix</i> .....	14
Tabel 4.1 Data Penjualan Hijab .....	22
Tabel 4.2 Hasil Label Pemrograman.....	24
Tabel 4.3 Transformasi Data.....	25
Tabel 4.4 Pelabelan Diskrit Label Kelas.....	26
Tabel 4.5 Nilai <i>Entropy</i> dan <i>Gain</i> .....	34
Tabel 4.6 <i>Confusion Matrix Data Testing</i> .....	40
Tabel 4.7 Hasil Uji Akurasi .....	40

## DAFTAR GAMBAR

Gambar 2.1 Semua Kemungkinan Penentuan Atribut .....	10
Gambar 2.2 Alternatif Penentuan Atribut dengan Nilai <i>Gain</i> .....	13
Gambar 2.3 <i>Decision Tree</i> untuk Data Cuaca.....	14
Gambar 2.4 Kurva AUC .....	16
Gambar 4.1 <i>Decision Tree</i> Penjualan Hijab.....	36
Gambar 4.2 Akar Pertama <i>Decision Tree</i> .....	37
Gambar 4.3 Nilai <i>True Decision Tree</i> .....	37
Gambar 4.4 Nilai <i>True</i> dan <i>False</i> Atribut Bintang .....	38
Gambar 4.5 Nilai <i>True</i> dan <i>False</i> Atribut <i>Brand</i> .....	38
Gambar 4.6 Nilai <i>True</i> dan <i>False</i> Atribut Harga .....	39
Gambar 4.7 Nilai <i>False Decision Tree</i> .....	39

## ABSTRAK

Husna, Faridatul. 2021. **Implementasi Data Mining Menggunakan Algoritma C4.5 pada Klasifikasi Penjualan Hijab**. Skripsi. Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing : (I) Dr. Hairur Rahman, M.Si. (II) Juhari, M.Si.

**Kata kunci** : akurasi, algoritma c4.5, *decision tree*, klasifikasi.

Klasifikasi merupakan suatu proses untuk menemukan properti-properti yang sama dalam suatu himpunan data untuk diklasifikasikan ke dalam kelas-kelas yang berbeda. Salah satu metode klasifikasi adalah *Decision Tree* menggunakan Algoritma C4.5. penelitian ini bertujuan untuk mengetahui model dan keakuratan algoritma dalam melakukan klasifikasi terhadap penjualan hijab dari beberapa *brand* hijab. Diperoleh model *Decision Tree* menggunakan Algoritma C4.5 dengan akar pertama adalah atribut Harga dimana yang menjadi akar pertama adalah atribut yang paling mempengaruhi penjualan hijab. Hasil perhitungan nilai akurasi adalah 87% sehingga model *Decision Tree* dan proses klasifikasi menggunakan Algoritma C4.5 tergolong baik.

## ABSTRACT

Husna, Faridatul. 2021. **Implementation of Data Mining Using the C4.5 Algorithm On the Hijab Sales Classification**. Thesis. Mathematics Study Program. Faculty of Science and Technology, Maulana Malik Ibrahim State Islamic University Malang. Advisors: (I) Dr. Hairur Rahman, M.Si. (II) Juhari, M.Si.

**Keywords:** accuracy, c4.5 algorithm, decision tree, classification.

A classification is a process to find the same properties in a data set to be classified into different classes. One of the classification methods is the Decision Tree using the C4.5 Algorithm. This research aims to determine the model and the accuracy of the C4.5 algorithm in classifying hijab sales from several hijab brands. The Decision Tree model is obtained using the C4.5 algorithm with the first root being the price attribute, where the first root is the attribute that most affected the sale of the hijab. The result of calculating the accuracy value is 87% so that the Decision Tree model and the classification process using the C4.5 Algorithm are classified as good.

## مخلص

الحسنى، فريدة. ٢٠٢١. تطبيق التنقيب عن البيانات باستخدام خوارزمية ج ٥.٤ (Algoritma C4.5) في تصنيف (Klasifikasi) مبيعات الحمار. البحث العلمي. برنامج دراسة الرياضيات، كلية العلوم والتكنولوجيا، جامعة مولانا مالك إبراهيم الإسلامية الحكومية، مالانج. مشرف: (١) الدكتور خير الحمن الماجستير (٢) الاستاذ جوهاري الماجستير.

الكلمات الرئيسية: الدقة (Akurasi)، الخوارزمية ج ٥.٤ (Algoritma C4.5)، شجرة القرار (Tree Decision)، التصنيف (Klasifikasi).

التصنيف (Klasifikasi) هو عملية للعثور على نفس الخصائص في مجموعة البيانات لتصنيفها إلى فئات مختلفة. وأما إحدى طريقة التصنيف (Klasifikasi) شجرة القرار (Decision Tree) باستخدام خوارزمية ج ٥.٤ (Algoritma C4.5). تهدف هذه الدراسة إلى تحديد النموذج ودقة الخوارزمية في التصنيف مبيعات الحمار من عدة براند الحمار. حصلوا على نموذج شجرة القرار (Decision Tree) باستخدام الخوارزمية ج ٥.٤ (Algoritma C4.5) مع اعتبار الجذر الأول هو سمت السعر حيث يكون الجذر الأول هو الذي سمت الأكثر تأثيراً المبيعات الحمار. و نتيجة حساب قيمة الدقة ٨٧٪ بحيث يستخدم نموذج شجرة القرار (Decision Tree) وعملية التصنيف الخوارزمية ج ٥.٤ (Algoritma C4.5) على أنها الأفضل التصنيف (Klasifikasi).



# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Indonesia mempunyai 16 subsektor kreatif yang menjadi fokus pengembangan, di antaranya yaitu musik, kuliner, *fashion*, kriya, fotografi, film, animasi video, aplikasi dan pengembangan gambar arsitektur desain interior, desain produk, desain komunikasi visual, periklanan, penerbitan, televisi, radio, seni rupa, dan seni pertunjukan. Berdasarkan data hasil survei yang dikeluarkan oleh BEKRAF (Badan Ekonomi Kreatif) pada tahun 2016, *fashion* menempati urutan kedua dengan presentase 18,01% yang memberikan kontribusi besar dalam pertumbuhan ekonomi kreatif di Indonesia. Wishnutama Kusubandio (Menteri Pariwisata dan Ekonomi Kreatif) mengatakan, bahwa pada tahun 2020 *fashion* mengalami peningkatan kontribusi untuk ekonomi kreatif di Indonesia yakni menempati urutan pertama (Yasyi, 2020). Sebagai negara yang dikenal dengan mayoritas penduduknya beragama Islam, tentu membuat kebutuhan sandang di Indonesia juga harus memperhatikan bagaimana kriteria pakaian umat Islam, salah satunya yaitu hijab. Perkembangan bisnis dalam dunia *fashion* khususnya hijab telah menjadi *trend setter* saat ini, faktanya dapat dilihat dari berbagai faktor seperti munculnya jenis-jenis model hijab, banyak komunitas hijabers yang bermunculan hingga diadakannya beragam bazar dan peragaan busana muslim.

Hingga saat ini, ada puluhan juta penduduk Indonesia yang telah mengenakan hijab, hal ini selaras dengan semakin berkembangnya industri pakaian muslim khususnya hijab. Hijab didesain sesuai *trend* dan kebutuhan muslimah saat ini,

bukan hanya untuk aktivitas formal, bahkan hijab juga didesain untuk muslimah yang gemar berolahraga. Cara pemasaran dan penjualan hijab kini tidak hanya secara konvensional namun juga dengan cara *online*. Secara konvensional maupun *online*, anjuran untuk melakukan jual beli telah disebutkan pada Firman Allah SWT:

*“Orang-orang yang memakan riba tidak dapat berdiri melainkan seperti berdirinya orang yang kemasukan setan karena gila. Yang demikian itu karena mereka berkata bahwa jual beli sama dengan riba. Padahal Allah telah menghalalkan jual beli dan mengharamkan riba. Barangsiapa mendapat peringatan dari Tuhannya, lalu dia berhenti, maka apa yang telah diperolehnyadahulu menjadi miliknya dan urusannya (terserah) kepada Allah. Barangsiapa mengulangi, maka mereka itu penghuni neraka, mereka kekal di dalamnya.” (Q.S Al-Baqarah: 275).*

Ayat tersebut telah menjelaskan bahwa Allah SWT menghalalkan untuk melakukan proses jual beli, tentunya jual beli yang dimaksudkan adalah jual beli yang baik. Ayat tersebut juga menjelaskan bahwa Allah SWT mengharamkan riba, oleh karena itu dalam transaksi jual beli terdapat hukum dan syarat-syarat jual beli. Tingginya angka pengguna hijab, membuat produsen hijab terus berinovasi untuk menarik perhatian konsumen. Belum diketahuinya informasi dari banyaknya data penjualan hijab, membuat produsen kurang maksimal dalam mengetahui daya minat dan pelayanannya terhadap konsumen.

Pertumbuhan yang pesat dari himpunan data sudah menciptakan kondisi dimana sekarang ini terdapat banyak data namun minim informasi, maka diperlukan proses klasifikasi penjualan hijab untuk mengolah data tersebut agar didapatkan

informasinya. Informasi dari data dapat ditemukan *data mining*. *Data mining* dapat diartikan sebagai temuan informasi baru dengan mencari suatu pola atau aturan tertentu dari suatu data dengan jumlah besar yang diharapkan dapat mengatasi kondisi kaya akan data namun minim informasi (Yudistiro, 2015). Salah satu teknik *data mining* yaitu klasifikasi sesuai untuk diterapkan pada kasus yang akan digunakan pada penelitian ini. Klasifikasi dilakukan dengan menggunakan data *set* yang dibandingkan untuk mengembangkan model yang mampu mengklasifikasikan seluruh data yang ada (Ramageri, 2010). Terdapat banyak model klasifikasi salah satunya adalah *decision tree* menggunakan algoritma C4.5. Seperti pada penelitian yang telah dilakukan oleh Lakshmi, dkk yang mengklasifikasikan risiko selama kehamilan menggunakan algoritma C4.5 dan penelitian yang dilakukan oleh Audio Alief Kautsar Hartama yang mengklasifikasikan penyakit hipertensi di RSUD Provinsi NTB hingga didapatkan informasi bahwa atribut sistolik merupakan atribut yang paling berpengaruh dalam proses klasifikasi penyakit hipertensi. Penelitian ini menggunakan algoritma C4.5 untuk mengklasifikasikan data penjualan hijab dari beberapa *brand* hijab dan menggunakan beberapa atribut.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang tersebut, rumusan masalah yang diangkat pada penelitian ini adalah sebagai berikut:

1. Bagaimana *model decision tree* yang dibangun algoritma C4.5 untuk klasifikasi data penjualan beberapa *brand* hijab?
2. Bagaimana akurasi algoritma C4.5 untuk klasifikasi penjualan beberapa *brand* hijab?

### 1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah disebutkan, tujuan penelitian ini adalah sebagai berikut:

1. Mengetahui *model decision tree* yang dibangun algoritma C4.5 untuk mengklasifikasikan penjualan beberapa *brand* hijab.
2. Mengetahui akurasi algoritma C4.5 untuk mengklasifikasikan penjualan beberapa *brand* hijab.

### 1.4 Manfaat Penelitian

Manfaat yang dapat diperoleh dari penelitian ini yaitu:

1. Dapat digunakan sebagai acuan untuk mengetahui faktor yang mempengaruhi daya minat pengguna hijab.
2. Dapat digunakan sebagai tambahan wawasan keilmuan pada mata kuliah pilihan komputasi, yaitu *data mining* untuk mengklasifikasikan suatu data menggunakan algoritma C4.5.

### 1.5 Batasan Masalah

Agar penelitian ini tidak menyimpang, maka permasalahan harus diberikan batasan. Adapun batasan masalah dalam penelitian ini yaitu:

1. Data penjualan hijab yang didapatkan dari pengajuan kepada *brand* hijab Zealco.fa dan *marketplace* Shopee.
2. Data yang digunakan adalah data penjualan hijab sampai bulan Mei 2021.

## 1.6 Sistematika Penulisan

Penulis menggunakan sistematika penulisan yang terdiri dari 5 bab, yaitu:

### Bab I Pendahuluan

Pendahuluan berisikan latar belakang, rumusan masalah, tujuan penelitian, batasan masalah, manfaat penelitian, dan sistematika penulisan.

### Bab II Kajian Pustaka

Kajian pustaka menjelaskan teori-teori yang mendasari penyelesaian masalah, meliputi *data mining*, klasifikasi, *decision tree*, algoritma C4.5, dan kajian keislaman yang terkait dengan penelitian.

### Bab III Metode Penelitian

Metode penelitian menjelaskan langkah-langkah penelitian, yaitu jenis penelitian, data dan sumber data, metode pengumpulan data, dan tahap-tahap penelitian.

### Bab IV Pembahasan

Bab ini menjelaskan tentang langka-langkah dalam proses pembentukan *decision tree* menggunakan algoritma C4.5 serta pengujian akurasiya menggunakan *software* Python.

### Bab V Penutup

Penutup berisikan kesimpulan dari hasil penelitian yang diperoleh dan saran untuk pengembangan penelitian selanjutnya.

## BAB II

### KAJIAN PUSTAKA

#### 2.1 Decision Tree

Menurut (Micahel J. A. Berry, 2004), *decisison tree* merupakan struktur yang dapat digunakan untuk membagi fakta yang besar menjadi pohon keputusan yang mempresentasikan aturan, kemudian aturan tersebut dapat dengan mudah untuk diinterpretasi oleh manusia. Manfaat lain dari pohon keputusan adalah untuk mengeksplorasi data, mencari keterkaitan tersembunyi antara sejumlah variabel *input* dengan variabel target. Pohon keputusan biasanya menunjukkan data dalam bentuk tabel dengan atribut dan *record*. Atribut menunjukkan suatu parameter yang dibuat sebagai kriteria dalam pembentukan *tree*.

*Decision tree* digunakan untuk menyelesaikan kasus-kasus di mana *output*-nya bernilai diskrit. Atribut yang digunakan pada *decision tree* juga bernilai diskrit. Jika pada atribut ditemukan nilai numerik maka dapat ditentukan rentang nilai untuk mengubah nilai atribut numerik ke nilai diskrit. Meskipun banyak variasi model *decision tree* dengan tingkat kemampuan dan syarat yang berbeda, pada umumnya beberapa ciri kasus yang dapat diterapkan pada *decision tree* adalah (Santosa, 2007):

1. *Data/example* dinyatakan dengan pasangan atribut dan nilainya. Misalnya atribut suatu *example* adalah temperatur dan nilainya adalah dingin atau panas.

2. Label/*output* data bernilai diskrit. Nilai *output* bisa iya atau tidak, diterima atau ditolak. Namun, dalam kasus yang lain, bisa jadi *output* yang dihasilkan tidak hanya dua kelas.
3. Data mempunyai *missing value*. Misal untuk beberapa *example*, terdapat atribut yang nilainya tidak diketahui, namun dalam keadaan seperti ini *decision tree* masih dapat memberikan solusi yang baik.

## 2.2 Klasifikasi

Klasifikasi dapat didefinisikan sebagai *supervise learning* yang membutuhkan label dalam prosesnya untuk mengekstrak model yang digunakan untuk memprediksi suatu label (Han dan Kamber, 2012). Proses dalam klasifikasi adalah untuk menemukan properti-properti yang sama dalam suatu himpunan obyek pada suatu *database* kemudian diklasifikasikan ke dalam kelas-kelas yang berbeda sesuai dengan model klasifikasi yang dipilih. Proses klasifikasi bertujuan untuk mencari model dari *training set* yang memisahkan atribut ke dalam kategori atau kelas yang sesuai, kemudian model tersebut digunakan untuk klasifikasi atribut yang kelasnya belum diketahui sebelumnya (Azwanti, 2018).

## 2.3 Algoritma C4.5

Salah satu metode klasifikasi yang dipakai dalam penelitian ini adalah klasifikasi pohon keputusan menggunakan Algoritma C4.5. Algoritma pohon keputusan merupakan algoritma yang kuat, populer, berbasis logika, mudah ditafsirkan, dapat diterapkan secara luas pada beberapa kasus dalam *data mining*, dan merupakan algoritma siap pakai sehingga memberikan kinerja yang baik dan

cukup mudah dimengerti (Lakshmi dkk, 2016). Algoritma C4.5 dikembangkan oleh Ross Quinlan sebagai terusan dari Algoritma ID3.

Pohon keputusan dibangun untuk mengklasifikasikan data yang dirancang sebagai *input* pada algoritma yang terdiri dari beberapa objek dan atribut. Proses pembuatan pohon keputusan yang pertama adalah mengentropi masing-masing kelas dan atribut, kemudian penghitungan perolehan informasi (Mienye dkk, 2019). Informasi yang didapatkan dari pohon keputusan akan lebih mudah dipahami karena atribut dengan nilai *gain* tertinggi akan menjadi akar pohon yang kemudian diikuti cabang-cabangnya. *Gain* merupakan tingkat pengaruh suatu atribut terhadap keputusan atau ukuran efektifitas suatu variabel dalam mengklasifikasikan data (Lakshmi dkk, 2013).

Menurut (Eska, 2016), untuk membuat sebuah pohon keputusan harus melalui beberapa tahapan berikut ini:

1. Menyiapkan *data training*. Data yang digunakan adalah data yang sudah ada dan sudah dikelompokkan pada kelas tertentu.

Tabel 2.1 adalah contoh data cuaca dan keputusan bermain atau tidak (Santosa, 2007).

Tabel 2.1 Data Cuaca dan Keputusan untuk Main atau Tidak

Cuaca	Temperatur	Kelembaban	Angin	Main atau Tidak
$x_1$	$x_2$	$x_3$	$x_4$	$s$
Cerah	Panas	Tinggi	Kecil	Tidak
Cerah	Panas	Tinggi	Besar	Tidak



Lanjutan Tabel 2.1 Data Cuaca dan Keputusan untuk Main atau Tidak

Cuaca $x_1$	Temperatur $x_2$	Kelembaban $x_3$	Angin $x_4$	Main atau Tidak $s$
Hujan	Dingin	Normal	Besar	Tidak
Mendung	Dingin	Normal	Besar	Ya
Cerah	Sedang	Tinggi	Kecil	Tidak
Cerah	Dingin	Normal	Kecil	Ya
Hujan	Sedang	Normal	Kecil	Ya
Cerah	Sedang	Normal	Besar	Ya
Mendung	Sedang	Tinggi	Besar	Ya
Mendung	Panas	Normal	Kecil	Ya
Hujan	Sedang	Tinggi	Besar	Tidak
Mendung	Panas	Tinggi	Kecil	Ya
Hujan	Sedang	Tinggi	Kecil	Ya
Hujan	Dingin	Normal	Kecil	Ya

2. Kemudian menentukan akar dari pohon. Dilakukan dengan menghitung nilai *entropy* total terlebih dahulu kemudian menghitung nilai *entropy* dari setiap nilai atribut. *Entropy* adalah nilai informasi yang menyatakan ukuran ketidakpastian (*impurity*) atribut dari sekumpulan obyek data (Lakshmi dkk, 2013). Berikut persamaan untuk menentukan nilai *entropy* (Kusrini & Luthfi, 2009):

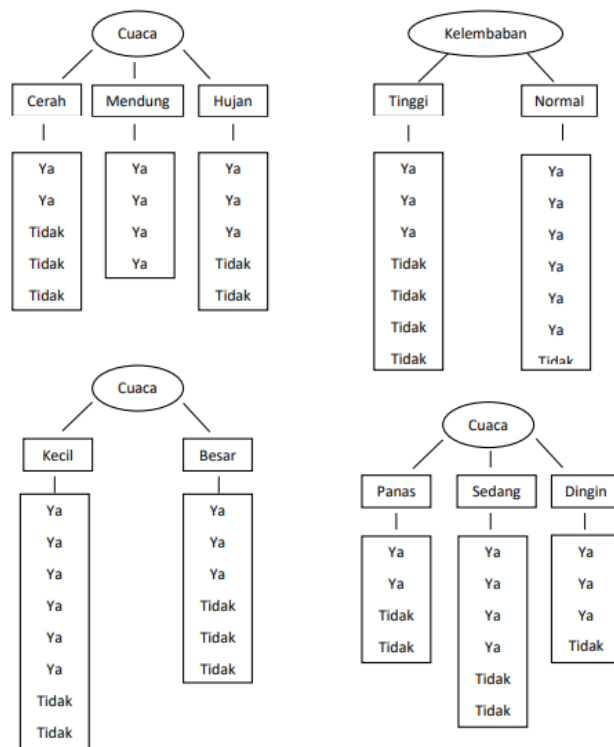
$$Entropy(s) = \sum_{i=1}^n -p_i \log_2 p_i \quad (2.1)$$

Keterangan:

$s$  = himpunan kasus

$n$  = jumlah partisi

$p_i$  = proporsi kelas  $i$  dalam data yang diproses pada node  $s$ . Nilai  $p_i$  didapat dengan membagikan jumlah baris data sebagai label kelas dengan jumlah baris semua data.



Gambar 2.1 Semua Kemungkinan Penentuan Atribut

Pengecekan dilakukan pada kelengkapan data cuaca dan keputusan untuk main atau tidak. Data sudah lengkap, tidak ada yang kurang, maka data tersebut dapat digunakan semuanya. Selanjutnya, akan dihitung nilai *entropy* total. Tabel 2.1, diketahui terdapat 9 keputusan main dan 5 keputusan tidak main. Maka dapat dihitung nilai *entropy* total dari *set output* data  $s$  yang berisikan 14 keputusan menggunakan persamaan (2.1).

$$entropy [9,5] = -\frac{9}{14} \log_2 \left( \frac{9}{14} \right) - \frac{5}{14} \log_2 \left( \frac{5}{14} \right) = 0.940$$

3. Setelah mencari nilai *entropy*, kemudian menghitung nilai *gain*. *Gain* merupakan tingkat pengaruh suatu atribut terhadap keputusan atau ukuran efektifitas suatu variabel dalam mengklasifikasikan data (Lakshmi dkk, 2013). maka dapat mengukur efektifitas suatu atribut dalam mengklasifikasikan data. Nilai *gain* dihitung menggunakan persamaan berikut (Kusrini & Luthfi, 2009):

$$Gain(s, A) = Entropy(s) - \sum_{i=1}^n \frac{|s_i|}{|s|} \cdot Entropy(s_i) \quad (2.2)$$

Keterangan:

$s$  = himpunan kasus

$A$  = nilai yang mungkin dari atribut  $A$

$n$  = jumlah partisi atribut  $A$

$|s_i|$  = subset dari  $s$  di mana  $A$  mempunyai nilai  $i$

$|s|$  = himpunan kasus

Dari data tersebut, kemudian dihitung nilai *entropy* untuk atribut cuaca.

- Cuaca = cerah

$$entropy [2,3] = -\frac{2}{5} \log_2 \left( \frac{2}{5} \right) - \frac{3}{5} \log_2 \left( \frac{3}{5} \right) = 0.971$$

- Cuaca = mendung

$$entropy [4,0] = -4 \log_2(4) - 0 \log_2(0) = 0$$

- Cuaca = hujan

$$entropy [3,2] = -\frac{3}{5} \log_2 \left( \frac{3}{5} \right) - \frac{2}{5} \log_2 \left( \frac{2}{5} \right) = 0.971$$

Maka *entropy* total untuk atribut cuaca adalah

$$\begin{aligned} \text{entropy}([2,3], [4,0], [3,2]) &= \frac{5}{14}(0.971) + \frac{4}{14}(0) + \frac{5}{14}(0.971) \\ &= 0.693 \end{aligned}$$

Selanjutnya akan dihitung nilai *gain*-nya, menggunakan persamaan (2.2). misal dipilih angin sebagai atribut  $A$ . Dari tabel 2.1 diketahui ada 9 keputusan ya dan 5 keputusan tidak. Maka dapat dihitung:

$$\text{nilai}(A) = [\text{kecil}, \text{besar}]$$

$$s = [9,5]$$

$$s_{\text{kecil}} \leftarrow [6,2]$$

$$s_{\text{besar}} \leftarrow [3,3]$$

Dihitung nilai *gain* angin (menggunakan persamaan 2.2) dengan mencari selisih sebelum dan sesudah dilakukan pemisahan.

$$\begin{aligned} \text{gain}(s, A) &= \text{entropy}(s) - \sum_{c \in \text{nilai}(A)} \frac{s_i}{s} \text{entropy}(s_i) \\ &= \text{entropy}(s) - \frac{8}{14} \text{entropy}(s_{\text{kecil}}) - \frac{6}{14} \text{entropy}(s_{\text{besar}}) \\ &= 0.940 - \left(\frac{8}{14}\right) 0.811 - \left(\frac{6}{14}\right) 1.0 \\ &= 0.048 \end{aligned}$$

Dari semua atribut, jika dihitung nilai *gain*-nya adalah:

- Angin = 0.048
- Cuaca = 0.247
- Temperatur = 0.029
- Kelembaban = 0.152

Hasil tersebut didapat dari nilai informasi *gain* sebelum dan sesudah pemecahan. Misalkan untuk atribut temperatur, informasi *gain* sebelum pemecahan adalah:

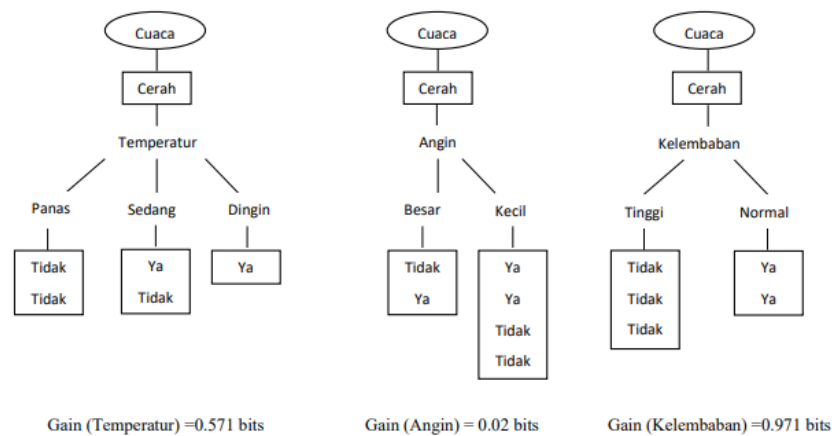
$$entropy [9,5] = -\frac{9}{14} \log_2 \left( \frac{9}{14} \right) - \frac{5}{14} \log_2 \left( \frac{5}{14} \right) = 0.940$$

Sedangkan nilai setelah pemecahan adalah:

$$entropy ([2,2], [4,2], [3,1]) = 0.911$$

Penentuan atribut untuk pemisahan data, digunakan nilai informasi *gain* yang paling besar.

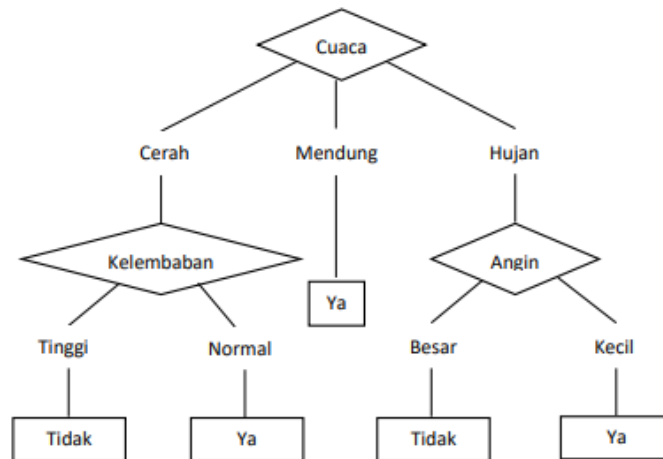
4. Kemudian mengulangi kembali langkah ke-2 hingga semua *record* terpartisi dengan baik.



Gambar 2.2 Alternatif Penentuan Atribut dengan Nilai *Gain*

5. Pemberhentian proses partisi pohon keputusan jika:
- Seluruh *record* pada simpul  $N$  mendapatkan kelas yang sama.
  - Tidak ada atribut yang dipartisi lagi.
  - Tidak ada *record* di dalam cabang yang kosong.

Nilai *entropy* dan *gain* yang telah didapatkan, maka akan terbentuk *decision tree* seperti pada Gambar 2.3.



Gambar 2.3 *Decision Tree* untuk Data Cuaca

## 2.4 Confusion Matrix

*Confusion matrix* adalah metode yang digunakan untuk melakukan perhitungan akurasi pada konsep *data mining*. *Confusion matrix* digambarkan dengan tabel yang menyatakan jumlah data uji yang benar dan jumlah data uji yang salah. Konsep ini dilakukan untuk mengevaluasi model klasifikasi berdasarkan proses hitung data *testing* yang memprediksi nilai benar atau salah. *Confusion matrix* terdiri atas data yang memiliki 2 kelas bersifat positif atau negatif. Tabel *confusion matrix* tersusun atas 4 sel yaitu *true positif*, *false positif*, *true negative*, dan *false negative* seperti tabel 2.2 (Rogers & Girolami, 2012).

Tabel 2.2 *Confusion Matrix*

Actual Class	Prediction Class	
	Kelas = Ya	Kelas = Tidak
Kelas = Ya	<i>True Positive (TP)</i>	<i>False Negative (FN)</i>
Kelas = Tidak	<i>False Positive (FP)</i>	<i>True Negative (TN)</i>

Berdasarkan Tabel 2.2 :

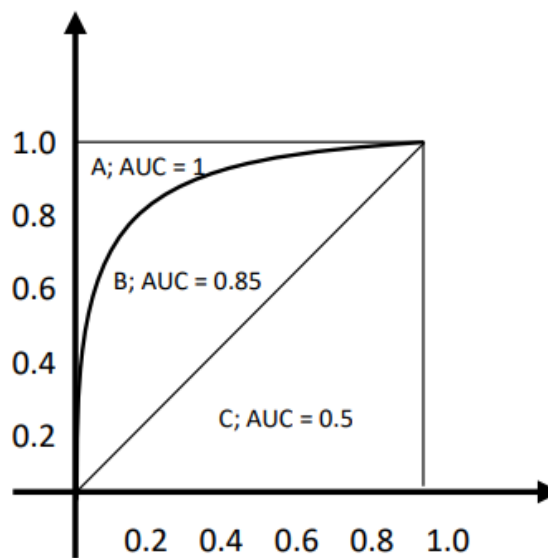
1. *True Positive* (TP) adalah jumlah *record* data positif yang diklasifikasikan sebagai nilai positif.
2. *False Positive* (FP) adalah jumlah *record* data negatif yang diklasifikasikan sebagai nilai positif.
3. *False Negative* (FN) adalah jumlah *record* data positif yang diklasifikasikan sebagai nilai negatif.
4. *True Negative* (TN) adalah jumlah *record* data negatif yang diklasifikasikan sebagai nilai positif.

Nilai yang dihasilkan melalui metode *confusion matrix* adalah berupa evaluasi akurasi sebagai berikut (Rahman dkk, 2017):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} 100\% \quad (2.3)$$

Pada akurasi klasifikasi data mining, nilai *Area Under Curve* (AUC) dapat dibagi menjadi beberapa kelompok yaitu (Florin, 2011):

1. 90% - 100% = Klasifikasi sangat baik
2. 80% - 90% = Klasifikasi baik
3. 70% - 80% = Klasifikasi cukup
4. 60% - 70% = Klasifikasi buruk
5. 50% - 60% = Klasifikasi salah



Gambar 2.4 Kurva AUC

## 2.5 K-Fold Cross Validation

Rata-rata keberhasilan suatu sistem dapat diketahui dengan melakukan *cross validation*, yaitu dengan melakukan iterasi dengan mengacak atribut *input* hingga sistem tersebut dapat diujikan untuk atribut *input* yang lain (Hartama, 2017). *K-fold cross validation* dilakukan dengan membagi data sejumlah *k-fold* yang diinginkan. Pada proses *cross validation*, data akan dikelompokkan dalam *k* buah partisi dengan ukuran yang sama, selanjutnya proses *testing* dan *training* dilakukan sebanyak *k* kali. Pada perulangan ke-*i* partisi akan menjadi data *testing* dan sisanya akan menjadi data *training*.

## 2.6 Kajian Keislaman

Islam sebagai agama yang sempurna, datang membawa keberkahan untuk mengatur segala aspek kehidupan manusia. Salah satu aspek kehidupan manusia yang diatur dalam Islam adalah proses kegiatan ekonomi. Kegiatan ekonomi dilakukan manusia untuk pemenuhan kebutuhan sandang, pangan, juga papan.



Ada banyak bentuk kegiatan ekonomi, salah satunya adalah proses jual beli. Dalam istilah fiqh, jual beli disebut sebagai *al-bai'* yang berarti menjual, mengganti, dan menukar sesuatu dengan sesuatu yang lain (Shobirin, 2015). Dalam Bahasa Arab, *al-bai'* biasanya digunakan sekaligus dengan pengertian lawannya. Maka, dapat dikatakan bahwa *al-bai'* tidak hanya berarti jual namun juga sekaligus beli. Secara istilah, jual beli dimaksudkan pada beberapa pengertian, yakni:

- a. Pelepasan hak milik dari yang satu kepada yang lain dengan perasaan rela, dapat dilakukan dengan tukar menukar barang dengan barang, ataupun barang dengan uang (Idris, 1986).
- b. Imam Taqiyuddin dalam kitab *Kifayatul al-Akhyar* menyebutkan bahwa jual beli merupakan proses saling tukar menukar harta, saling menerima, dapat dikelola (tasharruf) dengan ijab qobul, dengan cara yang sesuai dengan syariat.

Kegiatan jual beli telah dihalalkan oleh Allah, dengan ketentuan dan syariat-Nya yang telah dijelaskan pada Firman Allah pada surah Al-Baqarah ayat 275 yang artinya:

*“Orang-orang yang memakan riba tidak dapat berdiri melainkan seperti berdirinya orang yang kemasukan setan karena gila. Yang demikian itu karena mereka berkata bahwa jual beli sama dengan riba. Padahal Allah telah menghalalkan jual beli dan mengharamkan riba. Barangsiapa mendapat peringatan dari Tuhannya, lalu dia berhenti, maka apa yang telah diperolehnya dahulu menjadi miliknya dan urusannya (terserah) kepada Allah. Barangsiapa mengulangi, maka mereka itu penghuni neraka, mereka kekal di dalamnya.” (Q.S Al-Baqarah: 275).*

Dari ayat tersebut, dapat diketahui bahwa seorang muslim yang melakukan transaksi jual beli harus mengetahui dan sesuai dengan syarat-syarat praktik jual beli yang didasarkan pada ketentuan Al-Qur'an dan hadis, supaya dapat melaksanakannya sesuai dengan syariat sehingga tidak terjerumus ke dalam tindakan-tindakan yang dilarang dan diharamkan. Pada ringkasan Tafsir Ibnu Katsir (Ar-Rifa'i, 1999), dijelaskan bahwa para fuqaha mempersempit jalan yang dapat menembus riba, serta mereka mengharamkan jalan dan sarana yang mengantarkan pada riba, karena sesuatu yang mengantarkan pada yang haram adalah haram, sebagaimana wajibnya mengerjakan sesuatu yang menjadi prasyarat bagi terpenuhinya kewajiban yang lain. Anjuran untuk melakukan jual beli yang baik juga disebutkan pada sebuah hadis yang artinya, "*Sebaik-baik pekerjaan adalah pekerjaan dengan tangannya dan setiap jual beli yang dianggap baik.*"(HR. Ahmad, Al Bazzar, Ath Thobroni, dan selainnya).

## **BAB III**

### **METODE PENELITIAN**

#### **3.1 Sumber Data**

Jenis penelitian pada penelitian ini merupakan penelitian kuantitatif. Pada penelitian ini akan dilakukan analisis terhadap hasil berupa *decision tree* yang diperoleh dari subjek data yaitu data penjualan hijab dari bulan Maret 2020 sampai pada bulan Mei 2021.

Data yang digunakan adalah data penjualan hijab di website *marketplace* Shopee (alamat website terdapat pada lampiran 1) dan data penjualan hijab dari suatu *brand* Zealco.fa. Data yang didapatkan berupa data diskrit dan data numerik berdasarkan setiap atribut yang ditentukan. Faktor-faktor yang digunakan dalam penelitian ini adalah *brand* hijab, jenis hijab, bahan hijab, harga hijab, angka penjualan hijab, dan penilaian (bintang) yang diberikan konsumen pada pembelian produk hijab tersebut. Kemudian, kelas atau *output* yang telah ditentukan dalam penelitian ini yaitu daya minat konsumen terhadap hijab dengan label tinggi dan rendah.

#### **3.2 Variabel Penelitian**

Penelitian ini menggunakan dua variabel, yaitu variabel *independent* (bebas) dengan lima atribut yaitu *Brand*, Jenis, Bahan, Harga, dan Bintang. Dan variabel *dependent* (terikat) yaitu label/ kelas output pada penelitian ini yaitu daya minat konsumen berdasarkan angka terjual hijab dengan label rendah dan tinggi.

### 3.3 Langkah-Langkah Penelitian

Langkah pertama pada penelitian ini adalah menghitung nilai *entropy* total dari seluruh data menggunakan persamaan (2.1).

1. Kemudian menghitung nilai *entropy* dari setiap nilai pada atribut menggunakan persamaan (2.1).
2. Setelah nilai *entropy* dihitung, maka hitung nilai *gain* dari setiap atribut menggunakan persamaan (2.2).
3. Membuat akar *decision tree* menggunakan atribut yang memiliki nilai *gain* tertinggi.
4. Atribut yang memiliki nilai *gain* tertinggi kemudian dihitung nilai *entropy* totalnya. Misalnya pada atribut harga, terdapat dua nilai yaitu harga murah dan harga mahal. Maka dihitung nilai *entropy* total untuk kasus harga murah dan harga mahal menggunakan persamaan (2.2). Setelah itu dihitung nilai *gain* dari setiap atribut dengan kasus harga murah dan harga mahal menggunakan persamaan (2.2). Maka akan didapatkan nilai *gain* tertinggi dari kasus harga murah dan harga mahal dari perhitungan nilai *gain* tersebut yang kemudian dijadikan sebagai akar simpul pertama pada *decision tree*.
5. Melakukan proses pada poin ke-5 sampai semua atribut masuk pada *decision tree*.
6. Setelah *decision tree* terbentuk, maka dilakukan evaluasi akurasi untuk mengetahui keakuratan algoritma C4.5 menggunakan confusion matrix dan rumus *accuracy* pada persamaan (2.3).

7. Menguji *k-fold cross validation* untuk mengetahui akurasi dari keseluruhan proses *data mining*. Pengujian *k-fold cross validation* dilakukan dengan membagi *data testing* menjadi beberapa bagian.
8. Analisis hasil dari *decision tree* yang sudah terbentuk dan uji akurasi pada *confusion matrix*.

## BAB IV

### PEMBAHASAN

#### 4.1 Analisis Data

Pada bab ini akan dilakukan proses klasifikasi menggunakan algoritma C4.5 terhadap data penjualan beberapa *brand* hijab dengan bantuan pemrograman Google Colab (menggunakan bahasa Python). Google Colab adalah salah satu produk Google berbasis *cloud* yang dapat digunakan secara gratis yaitu *coding environment* bahasa pemrograman Python dengan format “*notebook*” (Adam, 2019)

##### 4.1.1 Pembersihan Data

Data yang digunakan berjumlah 104 data. Berikut adalah tabel data yang digunakan pada penelitian ini (Lampiran 3).

Tabel 4.1 Data Penjualan Hijab

<b>Brand</b>	<b>Jenis</b>	<b>Bahan</b>	<b>Harga</b>	<b>Bintang</b>	<b>Terjual</b>
Rabbani	Instan	Polyester	93.390	4,9	334 pcs
Zoya	Instan	Crepe	104.800	4,9	45 pcs
Nafisa	Instan Syari	Polyester	159.000	5	297 pcs
Elzatta	Segi Empat	Voal	130.000	4,8	49 pcs
Nadiraa	Instan	Crepe	89.000	4,8	1,1rb pcs
...	...	...	...	...	...
Ria Miranda	Segi Empat	Voal	325.000	5	161 pcs

Lanjutan Tabel 4.1 Data Penjualan Hijab

Ria Miranda	Segi Empat	Voal	325.000	4,9	120 pcs
Ria Miranda	Segi Empat	Voal	299.000	4,9	157 pcs
Ria Miranda	Segi Empat	Voal	195.000	5	50 pcs
Ria Miranda	Segi Empat	Voal	325.000	5	92 pcs
Ria Miranda	Segi Empat	Voal	325.000	5	92 pcs

Proses pembersihan data dilakukan untuk menghapus duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan penulisan (tipografi).

#### 4.1.2 Transformasi Data

Proses transformasi data dilakukan untuk mengubah atribut sesuai dengan format yang dapat diproses dalam program. Atribut *brand*, jenis, dan bahan dilakukan pengubahan nilai atribut dari diskrit menjadi numerik (label) pada program sehingga didapatkan label numerik pada Tabel 4.2.

Tabel 4.2 Hasil Label Pemrograman

Brand	Rabbani	0
	Zoya	1
	Nafisa	2
	Elzatta	3
	Nadiraa	4

	Zealcofa	5
	Kami	6
	Shafira	7
	Umama	8
	WMD	9
	Alsyahra	10
	Meccanism	11
	Ria Miranda	12
Jenis	Instan	0
	Segi Empat	1
	Pashmina	2
	Instan Syari	3
Bahan	Polyester	0
	Voal	1
	Cerutty	2
	Crepe	3

Tabel 4.3 adalah tampilan data setelah transformasi data yaitu mengubah nilai atribut diskrit menjadi label nilai atribut numerik (lampiran 4).

Tabel 4.3 Transformasi Data

<b>Brand</b>	<b>Jenis</b>	<b>Bahan</b>	<b>Harga</b>	<b>Bintang</b>	<b>Terjual</b>
0	0	0	93.390	4,9	334 pcs
1	0	3	104.800	4,9	45 pcs



2	3	0	159.000	5	297 pcs
3	1	1	130.000	4,8	49 pcs
4	0	3	89.000	4,8	1,1rb pcs
...	...	...	...	...	...
12	1	1	325.000	5	161 pcs
12	1	1	325.000	5	92 pcs
12	1	1	325.000	4,9	120 pcs
12	1	1	299.000	4,9	157 pcs
12	1	1	195.000	5	50 pcs

Selain atribut, kelas/ *label output* yang digunakan juga dilakukan perubahan nilai numerik ke nilai diskrit dengan mencari nilai rata-rata dari total penjualan pada data tersebut.

$$\text{rata - rata} = \frac{\text{total terjual}}{\text{banyak data}}$$

$$\text{rata - rata} = \frac{87445}{104} = 840,8$$

Setelah didapatkan nilai rata-rata dari data tersebut, ketentuan untuk perubahan nilai numerik ke diskrit adalah sebagai berikut:

$\text{terjual} \leq 841 \text{ pcs}$  : Rendah

$\text{terjual} > 841 \text{ pcs}$  : Tinggi

Tabel 4.4 adalah data yang telah dilakukan perubahan nilai label kelas numerik menjadi nilai label kelas diskrit (data selengkapnya terdapat pada lampiran 5).

Tabel 4.4 Pelabelan Diskrit Label Kelas

Brand	Jenis	Bahan	Harga	Bintang	Terjual
0	0	0	93.390	4,9	Rendah
1	0	3	104.800	4,9	Rendah
2	3	0	159.000	5	Rendah
3	1	1	130.000	4,8	Rendah
4	0	3	89.000	4,8	Tinggi
...	...	...	...	...	...
12	1	1	325.000	5	Rendah
12	1	1	325.000	5	Rendah
12	1	1	325.000	4,9	Rendah
12	1	1	299.000	4,9	Rendah
12	1	1	195.000	5	Rendah

#### 4.1.3 Proses Pembentukan *Decision Tree* Menggunakan Algoritma C4.5

Menghitung nilai *entropy* total, nilai *entropy* dari nilai setiap atribut kemudian dihitung nilai *gain* dari setiap atribut. Nilai *entropy* total dan *entropy* dari nilai setiap atribut pada data di lampiran ketiga dihitung menggunakan persamaan (2.1).

Diketahui jumlah kasusnya adalah 104 kasus dengan kelas terjual tinggi sebanyak 31 dan terjual rendah sebanyak 73.

Maka nilai *entropy* total adalah:

$$\text{entropy total [31,73]} = -\frac{31}{104} \log_2 \left( \frac{31}{104} \right) - \frac{73}{104} \log_2 \left( \frac{73}{104} \right)$$

$$= 0,878927$$

Kemudian dihitung nilai *entropy* dari nilai setiap atribut dan nilai *gain* dari setiap atribut.

a. Atribut *Brand*

- *Brand Rabbani*

Jumlah kasus : 12

Kelas tinggi : 2

Kelas rendah : 10

$$\begin{aligned} \text{entropy} &= -\frac{2}{12} \log_2 \left( \frac{2}{12} \right) - \frac{10}{12} \log_2 \left( \frac{10}{12} \right) \\ &= 0,650022 \end{aligned}$$

- *Brand Zoya*

Jumlah kasus : 8

Kelas tinggi : 1

Kelas rendah : 7

$$\begin{aligned} \text{entropy} &= -\frac{1}{8} \log_2 \left( \frac{1}{8} \right) - \frac{7}{8} \log_2 \left( \frac{7}{8} \right) \\ &= 0,543564 \end{aligned}$$

- *Brand Nafisa*

Jumlah kasus : 8

Kelas tinggi : 4

Kelas rendah : 4

$$\begin{aligned} \text{entropy} &= -\frac{4}{8} \log_2 \left( \frac{4}{8} \right) - \frac{4}{8} \log_2 \left( \frac{4}{8} \right) \\ &= 1 \end{aligned}$$

- *Brand Elzatta*

Jumlah kasus : 11

Kelas tinggi : 8

Kelas rendah : 3

$$\begin{aligned} \text{entropy} &= -\frac{8}{11} \log_2 \left( \frac{8}{11} \right) - \frac{3}{11} \log_2 \left( \frac{3}{11} \right) \\ &= 0,845351 \end{aligned}$$

- *Brand Nadiraa*

Jumlah kasus : 7

Kelas tinggi : 3

Kelas rendah : 4

$$\begin{aligned} \text{entropy} &= -\frac{3}{7} \log_2 \left( \frac{3}{7} \right) - \frac{4}{7} \log_2 \left( \frac{4}{7} \right) \\ &= 0,985228 \end{aligned}$$

- *Brand Zealcofa*

Jumlah kasus : 9

Kelas tinggi : 5

Kelas rendah : 4

$$\begin{aligned} \text{entropy} &= -\frac{5}{9} \log_2 \left( \frac{5}{9} \right) - \frac{4}{9} \log_2 \left( \frac{4}{9} \right) \\ &= 0,991076 \end{aligned}$$

- *Brand Kami*

Jumlah kasus : 8

Kelas tinggi : 1

Kelas rendah : 7

$$\begin{aligned} \text{entropy} &= -\frac{1}{8} \log_2 \left( \frac{1}{8} \right) - \frac{7}{8} \log_2 \left( \frac{7}{8} \right) \\ &= 0,543564 \end{aligned}$$

- *Brand Shafira*

Jumlah kasus : 6

Kelas tinggi : 0

Kelas rendah : 6

$$\begin{aligned} \text{entropy} &= -\frac{0}{6} \log_2 \left( \frac{0}{6} \right) - \frac{6}{6} \log_2 \left( \frac{6}{6} \right) \\ &= 0 \end{aligned}$$

- *Brand Umama*

Jumlah kasus : 8

Kelas tinggi : 2

Kelas rendah : 6

$$\begin{aligned} \text{entropy} &= -\frac{2}{8} \log_2 \left( \frac{2}{8} \right) - \frac{6}{8} \log_2 \left( \frac{6}{8} \right) \\ &= 0,811278 \end{aligned}$$

- *Brand WMD*

Jumlah kasus : 7

Kelas tinggi : 4

Kelas rendah : 3

$$\begin{aligned} \text{entropy} &= -\frac{4}{7} \log_2 \left( \frac{4}{7} \right) - \frac{3}{7} \log_2 \left( \frac{3}{7} \right) \\ &= 0,985228 \end{aligned}$$

- *Brand Alsyahra*

Jumlah kasus : 5

Kelas tinggi : 2

Kelas rendah : 3

$$\begin{aligned} \text{entropy} &= -\frac{2}{5} \log_2 \left( \frac{2}{5} \right) - \frac{3}{5} \log_2 \left( \frac{3}{5} \right) \\ &= 0,721928 \end{aligned}$$

- *Brand Meccanism*

Jumlah kasus : 7

Kelas tinggi : 0

Kelas rendah : 7

$$\begin{aligned} \text{entropy} &= -\frac{0}{7} \log_2 \left( \frac{0}{7} \right) - \frac{7}{7} \log_2 \left( \frac{7}{7} \right) \\ &= 0 \end{aligned}$$

- *Brand Ria Miranda*

Jumlah kasus : 8

Kelas tinggi: 0

Kelas rendah : 8

$$\begin{aligned} \text{entropy} &= -\frac{0}{8} \log_2 \left( \frac{0}{8} \right) - \frac{8}{8} \log_2 \left( \frac{8}{8} \right) \\ &= 0 \end{aligned}$$

$$\begin{aligned} \text{Gain}(s, A) &= 0,878927 - \left( \left( \frac{12}{100} \text{ entropy Rabbani} \right) + \right. \\ &\left. \left( \frac{8}{100} \text{ entropy Zoya} \right) + \dots + \left( \frac{7}{100} \text{ entropy Ria Miranda} \right) \right) \\ &= 0,238457 \end{aligned}$$

b. Atribut Jenis

- *Jenis Instan*

Jumlah kasus : 25

Kelas tinggi : 12

Kelas rendah : 13

$$\begin{aligned} \text{entropy} &= -\frac{12}{25} \log_2 \left( \frac{12}{25} \right) - \frac{13}{25} \log_2 \left( \frac{13}{25} \right) \\ &= 0,998846 \end{aligned}$$

- Jenis Segi Empat

Jumlah kasus : 38

Kelas tinggi : 10

Kelas rendah : 28

$$\begin{aligned} \text{entropy} &= -\frac{10}{38} \log_2 \left( \frac{10}{38} \right) - \frac{28}{38} \log_2 \left( \frac{28}{38} \right) \\ &= 0,841852 \end{aligned}$$

- Jenis Pashmina

Jumlah kasus : 23

Kelas tinggi : 4

Kelas rendah : 19

$$\begin{aligned} \text{entropy} &= -\frac{4}{23} \log_2 \left( \frac{4}{23} \right) - \frac{19}{23} \log_2 \left( \frac{19}{23} \right) \\ &= 0,666587 \end{aligned}$$

- Jenis Instan Syari

Jumlah kasus : 18

Kelas tinggi : 6

Kelas rendah : 12

$$\begin{aligned} \text{entropy} &= -\frac{6}{18} \log_2 \left( \frac{6}{18} \right) - \frac{12}{18} \log_2 \left( \frac{12}{18} \right) \\ &= 0,852405 \end{aligned}$$

$$\begin{aligned}
 \text{Gain}(s, A) &= 0,878927 \\
 &\quad - \left( \left( \frac{23}{100} \text{ entropy Instan} \right) \right. \\
 &\quad \left. + \left( \frac{37}{100} \text{ entropy Segi Empat} \right) + \dots \right. \\
 &\quad \left. + \left( \frac{17}{100} \text{ entropy Instan Syari} \right) \right) \\
 &= 0,044367
 \end{aligned}$$

c. Atribut Bahan

- Bahan Polyester

Jumlah kasus : 23

Kelas tinggi : 7

Kelas rendah : 16

$$\begin{aligned}
 \text{entropy} &= -\frac{7}{23} \log_2 \left( \frac{7}{23} \right) - \frac{16}{23} \log_2 \left( \frac{16}{23} \right) \\
 &= 0,886541
 \end{aligned}$$

- Bahan Cerutti

Jumlah kasus : 18

Kelas tinggi : 4

Kelas rendah : 14

$$\begin{aligned}
 \text{entropy} &= -\frac{4}{18} \log_2 \left( \frac{4}{18} \right) - \frac{14}{18} \log_2 \left( \frac{14}{18} \right) \\
 &= 0,764205
 \end{aligned}$$

- Bahan Crepe

Jumlah kasus : 20

Kelas tinggi: 10



Kelas rendah : 10

$$\begin{aligned} \text{entropy} &= -\frac{10}{20} \log_2 \left( \frac{10}{20} \right) - \frac{10}{20} \log_2 \left( \frac{10}{20} \right) \\ &= 1 \end{aligned}$$

- Bahan Voal

Jumlah kasus : 43

Kelas tinggi : 11

Kelas rendah : 32

$$\begin{aligned} \text{entropy} &= -\frac{11}{43} \log_2 \left( \frac{11}{43} \right) - \frac{32}{43} \log_2 \left( \frac{32}{43} \right) \\ &= 0,820364 \end{aligned}$$

$$\begin{aligned} \text{Gain}(s, A) &= 0,878927 - \left( \left( \frac{23}{100} \text{entropy Polyester} \right) + \right. \\ &\left. \left( \frac{17}{100} \text{entropy Cerutty} \right) + \dots + \left( \frac{41}{100} \text{entropy Voal} \right) \right) \\ &= 0,020492 \end{aligned}$$

Dilakukan penghitungan nilai *entropy* sehingga didapatkan nilai *gain* dari semua atribut.

Tabel 4.5 Nilai *Entropy* dan *Gain*

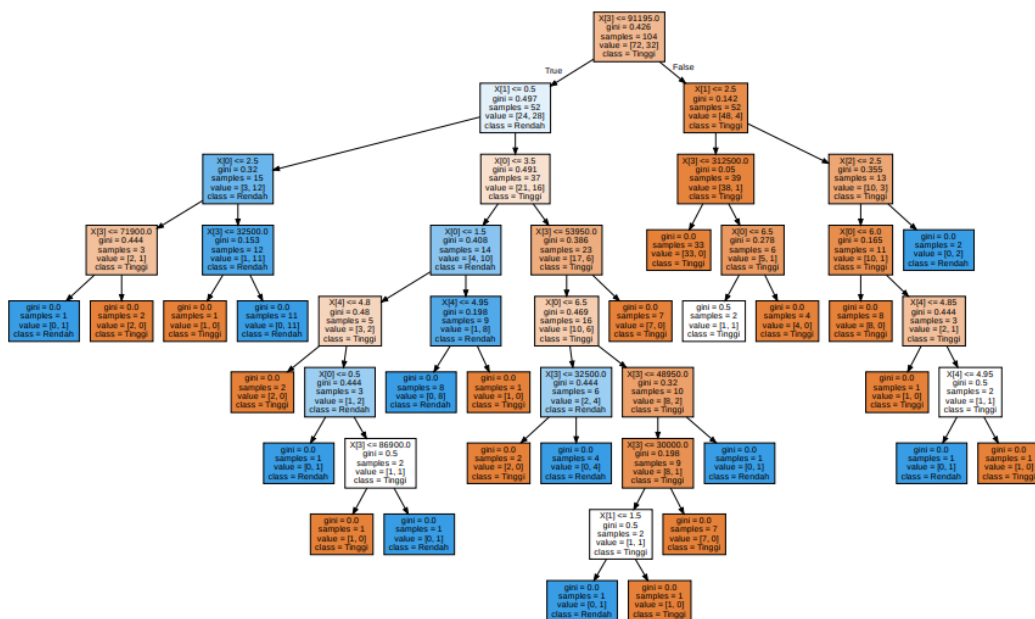
Atribut	Nilai	Jumlah Kasus	Kelas Tinggi	Kelas Rendah	<i>Entropy</i>	<i>Gain</i>
Total					0,878927	
<i>Brand</i>						0,238
	Rabbani	12	2	10	0,650022	

Lanjutan Tabel 4.5 Nilai *Entropy* dan *Gain*

Atribut	Nilai	Jumlah Kasus	Kelas Tinggi	Kelas Rendah	<i>Entropy</i>	<i>Gain</i>
	Zoya	8	1	7	0,543564	
	Nafisa	8	4	4	1	
	Elzatta	11	8	3	0,845351	
	Nadiraa	6	3	4	0,985228	
	Zealcofa	9	5	4	0,991076	
	Kami	8	1	7	0,543564	
	Shafira	6	0	6	0	
	Umama	7	2	6	0,811278	
	WMD	7	4	3	0,985228	
	Alsyahra	5	2	3	0,721928	
	Meccanism	7	0	7	0	
	RM	8	0	8	0	
Jenis						0,004
	Instan	25	12	13	0,987693	
	Segi Empat	37	10	27	0,841852	
	Pashmina	23	4	19	0,666587	
	Instan Syari	18	5	13	0,852405	
Bahan						0,025
	Polyester	23	7	16	0,886541	

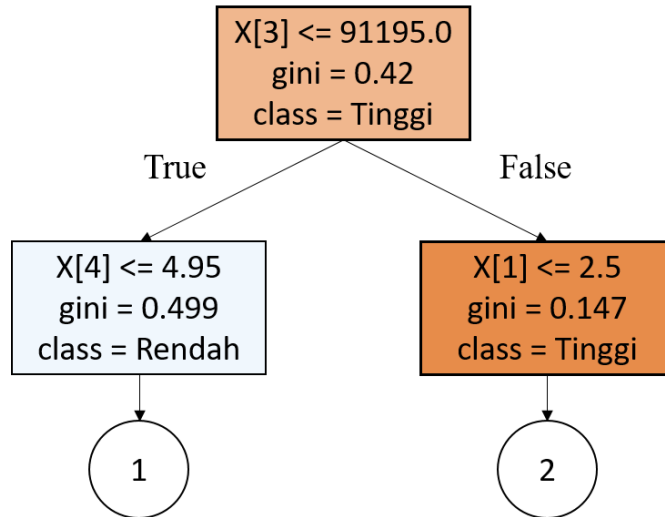
Berdasarkan nilai *gain* yang didapatkan dari semua atribut, nilai *gain* tertinggi adalah atribut Bintang, namun karena atribut yang memiliki nilai tertinggi adalah atribut Harga, maka yang menjadi akar pertama *decision tree* adalah atribut Harga.

Berikut adalah model *decision tree* yang telah dibangun dengan memproses data pada pemrograman Google Colab.

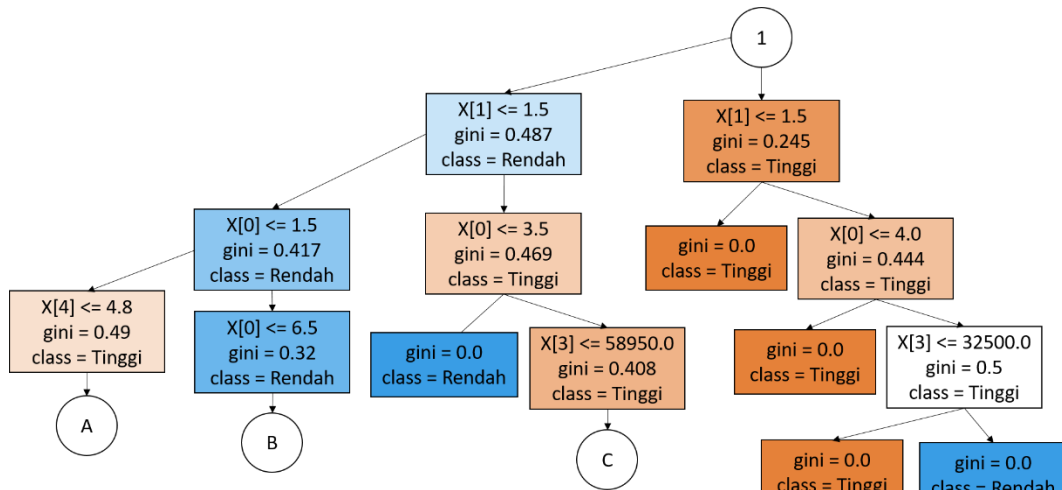


Gambar 4.1 *Decision Tree* Penjualan Hijab

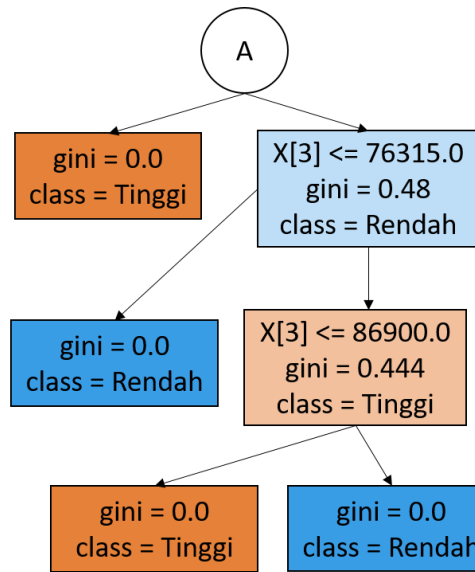
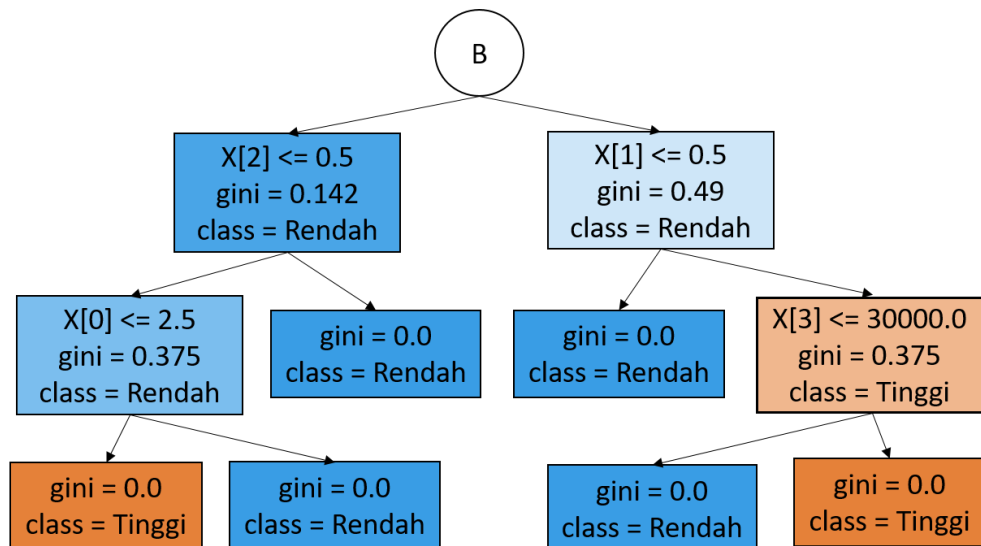
*Decision Tree* pada Gambar 4.1 secara lengkap dapat dilihat pada Gambar 4.2, 4.3, dst.

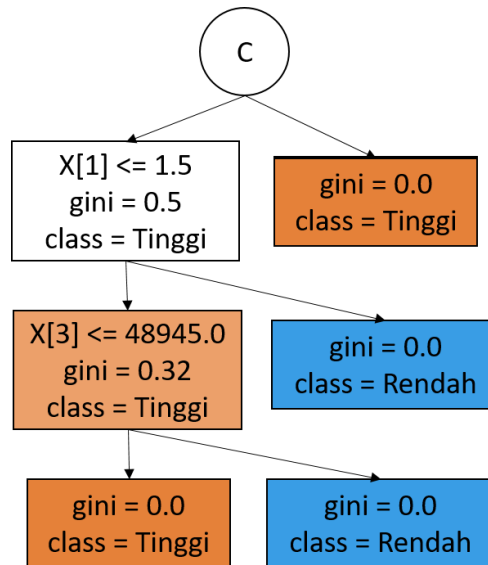


Gambar 4.2 Akar Pertama *Decision Tree*

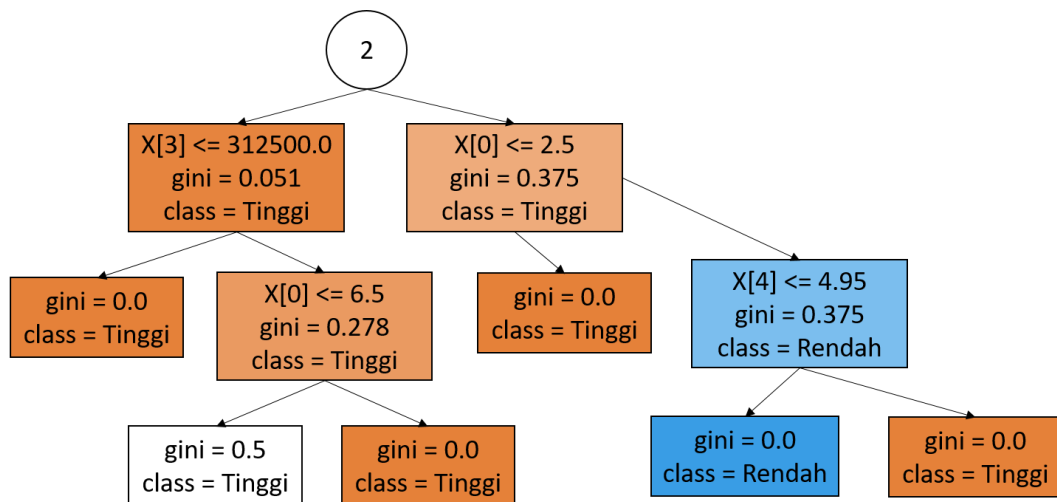


Gambar 4.3 Nilai True *Decision Tree*

Gambar 4.4 Nilai *True* dan *False* Atribut BintangGambar 4.5 Nilai *True* dan *False* Atribut Brand



Gambar 4.6 Nilai True dan False Atribut Harga



Gambar 4.7 Nilai False Decision Tree

*Decision Tree* pada Gambar 4.1 menghasilkan rules pada Lampiran 5.

## 4.2 Uji Akurasi Algoritma C4.5

Pengujian dilakukan untuk mengetahui kinerja metode algoritma C4.5 dalam melakukan klasifikasi terhadap kelas yang telah ditentukan. Pada pengujian ini, data yang berjumlah 104 data dibagi menjadi 70% data *training* dan 30% data *testing*.

Tabel 4.6 *Confusion Matrix Data Testing*

Kelas Observasi	Kelas Prediksi	
	Kelas = Tinggi	Kelas = Rendah
Kelas Sebenarnya		
Kelas = Tinggi	20	0
Kelas = Rendah	4	8

Tabel 4.6 adalah penghitungan *confussion matrix* dari data *testing* pada Google Colab (menggunakan bahasa Python). Jumlah *true positif* (tp) sebanyak 20 *record*, *false positif* (fp) sebanyak 4 *record*, *false negative* (fn) sebanyak 0 *record*, dan jumlah *true negative* (tn) sebanyak 8 *record*.

Setelah melakukan pengujian *confussion matrix*, data *testing* dilakukan uji akurasi, *precision*, dan *recall* pada Google Colab.

Tabel 4.7 Hasil Uji Akurasi

Uji	<i>Accuration</i>	<i>Precision</i>	<i>Recall</i>
Tingkat Akurasi	87%	83%	100%

Akurasi yang dihasilkan pada penghitungan menggunakan pemrograman Google Colab dihasilkan adalah :

$$\begin{aligned}
 \text{a. Akurasi} &= \frac{TP+TN}{TP+TN+FP+FN} 100\% \\
 &= \frac{20+8}{20+8+4+0} 100\% \\
 &= \frac{28}{32} 100\% \\
 &= 87\%
 \end{aligned}$$

$$\begin{aligned}
 \text{b. Precision} &= \frac{TP}{TP+FP} 100\% \\
 &= \frac{20}{20+4} 100\% \\
 &= \frac{20}{24} 100\% \\
 &= 83\%
 \end{aligned}$$

$$\begin{aligned}
 \text{c. Recall} &= \frac{TP}{TP+FN} 100\% \\
 &= \frac{20}{20+0} 100\% \\
 &= \frac{20}{20} 100\% \\
 &= 100\%
 \end{aligned}$$

### 4.3 K-Fold Cross Validation

Pengujian *cross validation* untuk *accuracy* dilakukan sebanyak 3 kali dengan *k-fold validation* yang berbeda pada operasi *cross validation*. Hasil pengujian adalah sebagai berikut :

1. *K-fold Validation 10* = 60%
2. *K-fold Validation 5* = 65%
3. *K-fold Validation 3* = 62%

### 4.4 Konsep Decision Tree untuk Mengatur Proses Jual Beli dalam Islam

Seperti yang telah disebutkan pada sub bab 2.7, bahwa “*Sebaik-baiknya pekerjaan adalah pekerjaan dengan tangannya dan setiap jual beli yang dianggap baik.*” (HR. Ahmad, Al Bazaar, Ath Thobroni, dan selainnya). Yang dimaksudkan jual beli yang baik adalah yang proses jual beli yang sesuai dengan syariat islam, tidak mengandung unsur riba dan kecurangan, juga memperjual belikan sesuatu



yang baik (bukan benda terlarang/ haram). Masa pandemi yang mengakibatkan sulitnya ekonomi, banyak orang terjun ke dalam usaha bisnis jual beli. Proses jual beli melalui *marketplace online* sangat meningkat saat pandemi. Persaingan yang semakin meningkat dapat mengakibatkan pelaku usaha melakukan segala cara termasuk mengesampingkan syariat islam agar produk yang dijual terjual dengan angka tinggi.

Dalam bisnis jual beli, terdapat suatu strategi yaitu *marketing*. Strategi ini digunakan untuk memenuhi kebutuhan dan kepuasan konsumen. Penggunaan *decision tree* dapat membantu pelaku usaha untuk memenuhi kebutuhan dan kepuasan konsumen, karena dari *decision tree* tersebut dapat diketahui faktor-faktor yang paling mempengaruhi daya tarik konsumen terhadap suatu produk. Oleh karena itu, tanpa melakukan kecurangan dalam proses jual beli, pelaku usaha dapat melakukan penjualan dengan angka yang tinggi karena mengetahui apa yang menjadi daya tarik dari konsumen.

Pelaku usaha juga tidak diperbolehkan menjual sesuatu yang terlarang/ haram. Dari hasil survei BEKRAF, disebutkan bahwa pada tahun 2020 *fashion* menempati urutan pertama dalam membantu pertumbuhan ekonomi kreatif di Indonesia. Dari hal tersebut, dan mengetahui bahwa Indonesia dikenal dengan penduduk mayoritas muslim, pelaku usaha dapat memulai usaha dibidang *fashion* terutama *muslim fashion* salah satunya hijab. Karena sebagai umat muslim, seorang muslimah diwajibkan untuk menutup aurat seperti yang diesbutkan dalam Firman Allah pada surah Al-Ahzab ayat 59 yang artinya:

*“Hai Nabi, katakanlah kepada istri-istrimu, anak-anak perempuanmu, dan istri-istri orang mukmin: “Hendaklah mereka menjulurkan hijabnya ke seluruh tubuh*

*mereka.” Yang demikian itu agar mereka lebih mudah untuk dikenali, sehingga mereka tidak diganggu. Dan Allah Maha Pengampun, Maha Penyayang.” (Q.S Al-Ahzab :59).*

Dari ayat tersebut, seorang pelaku usaha yang seharusnya melakukan jual beli yang baik, dapat menggunakan peluang untuk memenuhi kebutuhan sandang muslimah. Selain itu, pelaku usaha juga dapat menggunakan konsep *decision tree* dalam proses *marketing* jual beli.

## **BAB V**

### **PENUTUP**

#### **5.1 Kesimpulan**

Berdasarkan rumusan masalah beserta pembahasan di atas, dapat disimpulkan:

1. Dari *decision tree* yang terbentuk, didapatkan informasi bahwa faktor utama yang mempengaruhi penjualan hijab adalah harga, kemudian disusul oleh faktor *rate* bintang, jenis hijab, *brand* hijab, dan bahan hijab.
2. Hasil pengujian tingkat akurasi dengan Google Colab (menggunakan bahasa Python), diperoleh akurasi sebesar 87% menggunakan *confussion matrix*. Berdasarkan nilai Area Under Curve (AUC) yang telah disebutkan pada BAB II, maka klasifikasi algoritma C4.5 pada data yang digunakan tergolong baik.

#### **5.2 Saran**

Pada penelitian ini, penulis terfokus pada metode klasifikasi hanya dengan algoritma C4.5. Untuk meningkatkan kinerja dan menyempurnakan penelitian selanjutnya, penulis memberikan saran sebagai berikut :

1. Penelitian ini dapat dikembangkan dengan menggabungkan metode lain sehingga didapatkan hasil prediksi dari data.
2. Jumlah data ditambah, sehingga dapat diperoleh hasil akurasi fungsi algoritma yang lebih baik.
3. Pembuatan *user interface* agar tidak perlu menuliskan kembali kodingan, namun hanya perlu *input* data.

## DAFTAR PUSTAKA

- Adam, Rian. "Mengenal Google Colab". <https://structilmy.com/2019/05/mengenal-google-colab/>, diakses pada 12 Desember 2021 pukul 16.40.
- Ar-Rifa'i, M. N. 1999.. *Ringkasan Tafsir Ibnu Katsir*. Gema Insani.
- Azwanti, N. 2018. Analisa Algoritma C4.5 untuk Memprediksi Penjualan Motor pada PT. Capella Dinamik Nusantara Cabang Muka Kuning. *Informatika Mulawarman: Jurnal Ilmiah Ilmu Komputer*, 13(1), 6. <http://e-journals.unmul.ac.id/index.php/JIM/article/view/629>
- Eska, J. 2016. Penerapan Data Mining untuk Prediksi Penjualan Wallpaper Menggunakan Algoritma C4.5. *JURTEKSI (Jurnal Teknologi Dan Sistem Informasi)*, 2(2), 9–13. <https://osf.io/preprints/inarxiv/x6svc/>
- Gorunescu, Florin. 2011. *Data Mining: Concepts, Models, and Techniques*. Verlag Berlin Heidelberg: Springer .
- Hartama, A. A. K. 2017. *Klasifikasi Penyakit Hipertensi Menggunakan Algoritma C4.5 Studi Kasus RSUD Provinsi NTB*. 1–154.
- Han, J, Kamber, M, & Pei, J. 2012. *Data Mining: Concept and Techniques, Third Edition*. Waltham: Morgan Kaufmann Publishers.
- Kusrini & Luthfi, Emha T. 2009. *Algoritma Data Mining*. ANDI OFFSET.
- Lakshmi, B. N., Indumathi, T. S., & Ravi, N. (2016). A Study on C.5 Decision Tree Classification Algorithm for Risk Predictions During Pregnancy. *Procedia Technology*, 24, 1542–1549. <https://doi.org/10.1016/j.protcy.2016.05.128>
- Micahel J. A. Berry, G. L. 2004. *Data Mining Techniques For Marketing, Sales, and Customer Relationship Management*. Wiley Publishing.
- Mienye, I. D., Sun, Y., & Wang, Z. 2019. Prediction performance of improved decision tree-based algorithms: A review. *Procedia Manufacturing*, 35, 698–703. <https://doi.org/10.1016/j.promfg.2019.06.011>
- Ramageri, Bharati M. 2010. *Data Mining Techniques and Applications*. *Indian Journal of Computer Science and Engineering*. 1(4).
- Rahman, M. F., Alamsah, D., Darmawidjadja, M. I., & Nurma, I. 2017. Klasifikasi Untuk Diagnosa Diabetes Menggunakan Metode Bayesian Regularization

Neural Network (RBNN). *Jurnal Informatika*, 11(1), 36.  
<https://doi.org/10.26555/jifo.v11i1.a5452>

Rogers, Simon & Girolami, Mark. 2012. *A First Course in Machine Learning*. CRC Press Taylor & Francis Group.

Santosa, B. 2007. *Data Mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis* (1st ed.). Graha Ilmu.

Shobirin. 2015. Jual Beli dalam Pandangan Islam. *Jurnal IAIN Kudus*, 3(2).

Yasyi, Dini N. “Tahun 2020, Sektor Ekonomi Kreatif Akan Sumbang Rp.1.100 Triliun ke PDB Indonesia”, <https://www.goodnewsfromindonesia.id/2020/08/31/tahun-2020-sektor-ekonomi-kreatif-akan-sumbang-rp1-100-triliun-ke-pdb-indonesia>, diakses pada 2 Juni 2021 pukul 14.02.

Yudistiro, Muhammad R. 2015. PEMILIHAN BEASISWA BAGI MAHASISWA *Universias Diponegoro Journal*.

## LAMPIRAN

### Lampiran 1. Alamat Web Sumber Data

Rabbani	: <a href="https://shopee.co.id/rabbani.official?v=70f&amp;smtt=0.0.3">https://shopee.co.id/rabbani.official?v=70f&amp;smtt=0.0.3</a>
Zoya	: <a href="https://shopee.co.id/zoyaplaza.official?v=b3b&amp;smtt=0.0.3">https://shopee.co.id/zoyaplaza.official?v=b3b&amp;smtt=0.0.3</a>
Nafisa	: <a href="https://shopee.co.id/nafisaofficial?v=69d&amp;smtt=0.0.3">https://shopee.co.id/nafisaofficial?v=69d&amp;smtt=0.0.3</a>
Elzatta	: <a href="https://shopee.co.id/elhijab.official?v=900&amp;smtt=0.0.3">https://shopee.co.id/elhijab.official?v=900&amp;smtt=0.0.3</a>
Nadiraa	: <a href="https://shopee.co.id/nadiraa_hijab?v=271&amp;smtt=0.0.3">https://shopee.co.id/nadiraa_hijab?v=271&amp;smtt=0.0.3</a>
Kami	: <a href="https://shopee.co.id/kamiidea?v=fe5&amp;smtt=0.0.3">https://shopee.co.id/kamiidea?v=fe5&amp;smtt=0.0.3</a>
Shafira	: <a href="https://shopee.co.id/shafira.official?v=31b&amp;smtt=0.0.3">https://shopee.co.id/shafira.official?v=31b&amp;smtt=0.0.3</a>
Umama	: <a href="https://shopee.co.id/umamaofficial?v=4c6&amp;smtt=0.0.3">https://shopee.co.id/umamaofficial?v=4c6&amp;smtt=0.0.3</a>
WMD	: <a href="https://shopee.co.id/wmdfashion_official?v=cd7&amp;smtt=0.0.3">https://shopee.co.id/wmdfashion_official?v=cd7&amp;smtt=0.0.3</a>
Alsyaahra	: <a href="https://shopee.co.id/tokoalsyaahra?v=341&amp;smtt=0.0.3">https://shopee.co.id/tokoalsyaahra?v=341&amp;smtt=0.0.3</a>
Meccanism	: <a href="https://shopee.co.id/meccanismbyzaskiamecca.?v=568&amp;smtt=0.0.3">https://shopee.co.id/meccanismbyzaskiamecca.?v=568&amp;smtt=0.0.3</a>
Ria Miranda	: <a href="https://shopee.co.id/riamirandaofficial?v=d84&amp;smtt=0.0.3">https://shopee.co.id/riamirandaofficial?v=d84&amp;smtt=0.0.3</a>

## Lampiran 2. Script Pemrograman

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
from sklearn.impute import SimpleImputer
from sklearn import tree
import os
os.environ["PATH"] += os.pathsep + 'C:\Program Files\Graphviz\
bin'
import graphviz
from sklearn.tree import DecisionTreeClassifier
```

### Menampilkan data

```
dataset = pd.read_csv('dataskripsi24.csv', delimiter=';', head
er=0)
```

### Mengubah nilai diskrit menjadi label numerik

```
dataset['Brand'] = pd.factorize(dataset.Brand)[0]
dataset['Jenis'] = pd.factorize(dataset.Jenis)[0]
dataset['Bahan'] = pd.factorize(dataset.Bahan)[0]
dataset['Terjual'] = pd.factorize(dataset.Terjual)[0]
```

### Membangun *decision tree*

```
model = tree.DecisionTreeClassifier(random_state=0, max_depth=
None, min_samples_split=2, min_samples_leaf=1, min_weight_fract
ion_leaf=0, max_leaf_nodes=None, min_impurity_decrease=0)
```

```
namakelas = ['Tinggi', 'Rendah']
x = dataset.iloc[:, :-1].values
y = dataset.iloc[:, -1].values
```

```
features = x
target = y
```

```
clf = model.fit(features, target)
```

```
dot_data = tree.export_graphviz(model, out_file=None, class_na
mes=["Tinggi", "Rendah"], filled=True)
```

```
graph = graphviz.Source(dot_data)
graph.view()
```

## Menghitung Akurasi

```
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score, confusion_matrix,
classification_report
from sklearn.tree import DecisionTreeClassifier
from sklearn.model_selection import KFold
from sklearn.model_selection import cross_val_score
from numpy import mean
from numpy import std

atr_dataset = dataset.drop(columns = 'Terjual')
atr_dataset

cls_dataset = dataset['Terjual']
cls_dataset

xtrain, xtest, ytrain, ytest, = train_test_split(atr_dataset,
cls_dataset, test_size=0.3, random_state=0)
tree_dataset = DecisionTreeClassifier(random_state=0)
tree_dataset.fit(xtrain, ytrain)
cv = KFold(n_splits=3, random_state=0, shuffle=True)

scores = cross_val_score(model, xtest, ytest, scoring='accuracy',
cv=cv, n_jobs=-1)

print('Accuracy: %.3f (%.3f)' % (mean(scores), std(scores)))

y_pred = tree_dataset.predict(xtest)
cm = confusion_matrix(ytest, y_pred)
print("Confussion matrix = cm")
akurasi = classification_report(ytest, y_pred)
print("Tingkat Akurasi Algoritma C4.5")
print("Akurasi :", akurasi)
akurasi = accuracy_score(ytest, y_pred)
print("Tingkat Akurasi %d persen" %(akurasi*100))
```



**Lampiran 3. Data Penjualan Hijab**

<b>No</b>	<b>Brand</b>	<b>Jenis</b>	<b>Bahan</b>	<b>Harga</b>	<b>Bintang</b>	<b>Terjual</b>
1.	Rabbani	Instan	Polyester	93.390	4,9	334 pcs
2.	Rabbani	Segi Empat	Voal	67.830	4,9	1,5rb pcs
3.	Rabbani	Pashmina	Cerutty	99.800	5	90 pcs
4.	Rabbani	Instan	Crepe	168.900	4,8	73 pcs
5.	Rabbani	Segi Empat	Voal	82.800	4,7	85 pcs
6.	Rabbani	Instan	Crepe	204.800	4,6	236 pcs
7.	Rabbani	Instan	Polyester	104.800	4,9	82 pcs
8.	Rabbani	Instan	Polyester	64.800	4,9	4,1rb pcs
9.	Rabbani	Instan	Polyester	84.800	4,9	340 pcs
10.	Rabbani	Instan Syari	Polyester	126.000	4,9	245 pcs
11.	Rabbani	Instan Syari	Polyester	128.800	4,9	148 pcs
12.	Rabbani	Segi Empat	Voal	36.800	4,7	572 pcs
13.	Zoya	Instan	Crepe	104.800	4,9	45 pcs
14.	Zoya	Instan	Cerutty	124.800	4,9	90 pcs
15.	Zoya	Instan Syari	Cerutty	169.800	4,9	109 pcs
16.	Zoya	Instan	Crepe	124.800	4,9	71 pcs
17.	Zoya	Instan Syari	Cerutty	289.800	4,8	101 pcs
18.	Zoya	Segi Empat	Voal	84.800	4,9	18 pcs
19.	Zoya	Instan Syari	Cerutty	209.000	4,7	110 pc
20.	Zoya	Segi Empat	Voal	89.000	4,9	2,1rb pcs
21.	Nafisa	Instan Syari	Polyester	159.000	5	297 pcs
22.	Nafisa	Instan Syari	Polyester	159.000	4,9	435 pcs
23.	Nafisa	Segi Empat	Voal	89.000	4,9	2,2rb pcs
24.	Nafisa	Instan Syari	Polyester	179.000	4,8	141 pcs
25.	Nafisa	Segi Empat	Voal	89.000	4,6	1,4rb pcs
26.	Nafisa	Segi Empat	Voal	69.000	4,8	1,4rb pcs
27.	Nafisa	Segi Empat	Voal	89.000	4,9	1rb pcs
28.	Nafisa	Instan	Polyester	79.000	4,8	321 pcs
29.	Elzatta	Segi Empat	Voal	130.000	4,8	49 pc

30.	Elzatta	Segi Empat	Polyester	45.000	4,8	1,8rb pc
31.	Elzatta	Instan Syari	Polyester	89.000	4,9	1,1rb pcs
32.	Elzatta	Instan	Polyester	55.000	4,9	1,6rb pcs
33.	Elzatta	Instan Syari	Polyester	80.000	4,9	3,5rb pcs
34.	Elzatta	Segi Empat	Voal	25.900	4,8	2,9rb pcs
35.	Elzatta	Pashmina	Voal	129.000	4,8	39 pcs
36.	Elzatta	Pashmina	Polyester	84.000	5	17 pcs
37.	Elzatta	Instan	Polyester	69.000	4,9	1,7rb pcs
38.	Elzatta	Instan	Crepe	89.000	4,8	3,2rb pc
39.	Elzatta	Instan Syari	Crepe	189.000	4,9	1,4rb pc
40.	Nadiraa	Instan	Crepe	89.000	4,8	1,1rb pc
41.	Nadiraa	Pashmina	Cerutty	79.000	4,9	465 pc
42.	Nadiraa	Instan Syari	Crepe	79.000	4,9	169 pc
43.	Nadiraa	Pashmina	Voal	89.000	4,6	429 pc
44.	Nadiraa	Pashmina	Voal	89.000	4,7	152 pc
45.	Nadiraa	Instan	Crepe	49.000	4,8	2,7rb pc
46.	Nadiraa	Instan	Crepe	35.000	4,8	1,7rb pc
47.	Zealco.fa	Pashmina	Cerutty	40.000	4,9	3,5rb pc
48.	Zealco.fa	Instan	Crepe	38.710	4,8	3,2rb pc
49.	Zealco.fa	Pashmina	Cerutty	35.000	4,8	1,1rb pc
50.	Zealco.fa	Instan Syari	Crepe	35.000	4,7	1,8rb pc
51.	Zealco.fa	Pashmina	Cerutty	35.000	5	1,1rb pcs
52.	Zealco.fa	Pashmina	Crepe	30.000	5	325 pcs
53.	Zealco.fa	Segi Empat	Crepe	30.000	5	748 pcs
54.	Zealco.fa	Segi Empat	Voal	55.000	5	214 pcs
55.	Zealco.fa	Instan	Crepe	30.000	5	662 pcs
56.	Kami	Segi Empat	Voal	325.000	4,9	966 pcs
57.	Kami	Segi Empat	Voal	150.000	5	276 pcs
58.	Kami	Segi Empat	Voal	300.000	5	171 pcs
59.	Kami	Segi Empat	Voal	300.000	5	125 pcs
60.	Kami	Segi Empat	Voal	275.000	5	389 pcs

61.	Kami	Segi Empat	Voal	100.000	4,9	559 pcs
62.	Kami	Segi Empat	Voal	325.000	4,9	93 pcs
63.	Kami	Segi Empat	Voal	300.000	4,9	39 pcs
64.	Shafira	Pashmina	Voal	199.000	5	61 pcs
65.	Shafira	Pashmina	Polyester	199.000	4,9	44 pcs
66.	Shafira	Pashmina	Cerutty	199.000	5	9 pcs
67.	Shafira	Pashmina	Voal	199.000	5	44 pcs
68.	Shafira	Segi Empat	Voal	229.000	4,9	162 pcs
69.	Shafira	Segi Empat	Voal	349.000	4,8	5 pcs
70.	Umama	Segi Empat	Voal	25.000	4,9	4,4rb pcs
71.	Umama	Pashmina	Voal	35.000	4,9	600 pcs
72.	Umama	Pashmina	Voal	25.000	4,9	422 pcs
73.	Umama	Segi Empat	Voal	45.000	5	406 pcs
74.	Umama	Segi Empat	Polyester	40.000	4,9	163 pcs
75.	Umama	Pashmina	Polyester	35.000	4,8	450 pcs
76.	Umama	Instan	Cerutty	35.000	4,7	1,9rb pcs
77.	Umama	Segi Empat	Cerutty	45.000	4,9	102 pcs
78.	WMD	Instan	Crepe	39.900	4,8	4,4rb pcs
79.	WMD	Pashmina	Voal	44.990	4,9	436 pcs
80.	WMD	Pashmina	Voal	52.990	4,8	3,4rb pcs
81.	WMD	Instan	Crepe	49.500	4,8	6,3rb pcs
82.	WMD	Segi Empat	Voal	35.000	4,8	762 pcs
83.	WMD	Instan	Polyester	37.500	4,7	1,1rb pcs
84.	WMD	Instan Syari	Crepe	65.000	4,7	370 pcs
85.	Alsyahra	Instan Syari	Cerutty	210.000	4,9	905 pcs
86.	Alsyahra	Instan Syari	Cerutty	270.000	5	369 pcs
87.	Alsyahra	Instan Syari	Crepe	190.000	4,9	806 pcs
88.	Alsyahra	Pashmina	Cerutty	170.000	4,8	47 pcs
89.	Alsyahra	Instan Syari	Cerutty	140.000	4,8	337 pcs
90.	Meccanism	Instan	Polyester	105.000	4,9	485 pcs
91.	Meccanism	Instan	Polyester	105.000	4,9	547 pcs

92.	Meccanism	Pashmina	Cerutty	110.000	4,9	147 pcs
93.	Meccanism	Pashmina	Cerutty	85.000	4,9	65 pcs
94.	Meccanism	Segi Empat	Voal	129.900	4,9	22 pcs
95.	Meccanism	Instan	Crepe	110.000	5	124 pcs
96.	Meccanism	Instan	Polyester	99.000	5	384 pcs
97.	Ria Miranda	Segi Empat	Voal	215.000	4,9	92 pcs
98.	Ria Miranda	Pashmina	Cerutty	195.000	5	7 pcs
99.	Ria Miranda	Segi Empat	Voal	325.000	5	161 pcs
100.	Ria Miranda	Segi Empat	Voal	325.000	5	92 pcs
101.	Ria Miranda	Segi Empat	Voal	325.000	4,9	120 pcs
102.	Ria Miranda	Segi Empat	Voal	299.000	4,9	157 pcs
103.	Ria Miranda	Segi Empat	Voal	195.000	5	50 pcs
104.	Ria Miranda	Segi Empat	Voal	275.000	5	62 pcs

**Lampiran 4.** Data setelah Pembersihan Data

No	Brand	Jenis	Bahan	Harga	Bintang	Terjual
1.	0	0	0	93.390	4,9	334 pcs
2.	0	1	1	67.830	4,9	1,5rb pcs
3.	0	2	2	99.800	5	90 pcs
4.	0	0	3	168.900	4,8	73 pcs
5.	0	1	1	82.800	4,7	85 pcs
6.	0	0	3	204.800	4,6	236 pcs
7.	0	0	0	104.800	4,9	82 pcs
8.	0	0	0	64.800	4,9	4,1rb pcs
9.	0	0	0	84.800	4,9	340 pcs
10.	0	3	0	126.000	4,9	245 pcs
11.	0	3	0	128.800	4,9	148 pcs
12.	0	1	1	36.800	4,7	572 pcs
13.	1	0	3	104.800	4,9	45 pcs
14.	1	0	2	124.800	4,9	90 pcs

15.	1	3	2	169.800	4,9	109 pcs
16.	1	0	3	124.800	4,9	71 pcs
17.	1	3	2	289.800	4,8	101 pcs
18.	1	1	1	84.800	4,9	18 pcs
19.	1	3	2	209.000	4,7	110 pc
20.	1	1	1	89.000	4,9	2,1rb pcs
21.	2	3	0	159.000	5	297 pcs
22.	2	3	0	159.000	4,9	435 pcs
23.	2	1	1	89.000	4,9	2,2rb pcs
24.	2	3	0	179.000	4,8	141 pcs
25.	2	1	1	89.000	4,6	1,4rb pcs
26.	2	1	1	69.000	4,8	1,4rb pcs
27.	2	1	1	89.000	4,9	1rb pcs
28.	2	0	0	79.000	4,8	321 pcs
29.	3	1	1	130.000	4,8	49 pc
30.	3	1	0	45.000	4,8	1,8rb pc
31.	3	3	0	89.000	4,9	1,1rb pcs
32.	3	0	0	55.000	4,9	1,6rb pcs
33.	3	3	0	80.000	4,9	3,5rb pcs
34.	3	1	1	25.900	4,8	2,9rb pcs
35.	3	2	1	129.000	4,8	39 pcs
36.	3	2	0	84.000	5	17 pcs
37.	3	0	0	69.000	4,9	1,7rb pcs
38.	3	0	3	89.000	4,8	3,2rb pc
39.	3	3	3	189.000	4,9	1,4rb pc
40.	4	0	3	89.000	4,8	1,1rb pc
41.	4	2	2	79.000	4,9	465 pc
42.	4	3	3	79.000	4,9	169 pc
43.	4	2	1	89.000	4,6	429 pc
44.	4	2	1	89.000	4,7	152 pc
45.	4	0	3	35.000	4,8	1,7rb pc

46.	5	2	2	40.000	4,9	3,5rb pc
47.	5	0	3	38.710	4,8	3,2rb pc
48.	5	2	2	35.000	4,8	1,1rb pc
49.	5	3	3	35.000	4,7	1,8rb pc
50.	5	2	2	35.000	5	1,1rb pcs
51.	5	2	3	30.000	5	325 pcs
52.	5	1	3	30.000	5	748 pcs
53.	5	1	1	55.000	5	214 pcs
54.	5	0	3	30.000	5	662 pcs
55.	6	1	1	325.000	4,9	966 pcs
56.	6	1	1	150.000	5	276 pcs
57.	6	1	1	300.000	5	171 pcs
58.	6	1	1	300.000	5	125 pcs
59.	6	1	1	275.000	5	389 pcs
60.	6	1	1	100.000	4,9	559 pcs
61.	6	1	1	325.000	4,9	93 pcs
62.	6	1	1	300.000	4,9	39 pcs
63.	7	2	1	199.000	5	61 pcs
64.	7	2	0	199.000	4,9	44 pcs
65.	7	2	2	199.000	5	9 pcs
66.	7	2	1	199.000	5	44 pcs
67.	7	1	1	229.000	4,9	162 pcs
68.	7	1	1	349.000	4,8	5 pcs
69.	8	1	1	25.000	4,9	4,4rb pcs
70.	8	2	1	35.000	4,9	600 pcs
71.	8	2	1	25.000	4,9	422 pcs
72.	8	1	1	45.000	5	406 pcs
73.	8	1	0	40.000	4,9	163 pcs
74.	8	2	0	35.000	4,8	450 pcs
75.	8	1	2	45.000	4,9	102 pcs
76.	9	0	3	39.900	4,8	4,4rb pcs

77.	9	2	1	44.990	4,9	436 pcs
78.	9	2	1	52.990	4,8	3,4rb pcs
79.	9	0	3	49.500	4,8	6,3rb pcs
80.	9	1	1	35.000	4,8	762 pcs
81.	9	0	0	37.500	4,7	1,1rb pcs
82.	9	3	3	65.000	4,7	370 pcs
83.	10		2	210.000	4,9	905 pcs
84.	10	3	2	270.000	5	369 pcs
85.	10	3	3	190.000	4,9	806 pcs
86.	10	2	2	170.000	4,8	47 pcs
87.	11	0	0	105.000	4,9	485 pcs
88.	11	0	0	105.000	4,9	547 pcs
89.	11	2	2	110.000	4,9	147 pcs
90.	11	2	2	85.000	4,9	65 pcs
91.	11	1	1	129.900	4,9	22 pcs
92.	11	0	3	110.000	5	124 pcs
93.	11	0	0	99.000	5	384 pcs
94.	12	1	1	215.000	4,9	92 pcs
95.	12	2	2	195.000	5	7 pcs
96.	12	1	1	325.000	5	161 pcs
97.	12	1	1	325.000	5	92 pcs
98.	12	1	1	325.000	4,9	120 pcs
99.	12	1	1	299.000	4,9	157 pcs
100.	12	1	1	195.000	5	50 pcs

**Lampiran 5.** Data setelah Transformasi Data

<b>No</b>	<b>Brand</b>	<b>Jenis</b>	<b>Bahan</b>	<b>Harga</b>	<b>Bintang</b>	<b>Terjual</b>
1.	0	0	0	93.390	4,9	Rendah
2.	0	1	1	67.830	4,9	Tinggi
3.	0	2	2	99.800	5	Rendah
4.	0	0	3	168.900	4,8	Rendah
5.	0	1	1	82.800	4,7	Rendah
6.	0	0	3	204.800	4,6	Rendah
7.	0	0	0	104.800	4,9	Rendah
8.	0	0	0	64.800	4,9	Tinggi
9.	0	0	0	84.800	4,9	Rendah
10.	0	3	0	126.000	4,9	Rendah
11.	0	3	0	128.800	4,9	Rendah
12.	0	1	1	36.800	4,7	Rendah
13.	1	0	3	104.800	4,9	Rendah
14.	1	0	2	124.800	4,9	Rendah
15.	1	3	2	169.800	4,9	Rendah
16.	1	0	3	124.800	4,9	Rendah
17.	1	3	2	289.800	4,8	Rendah
18.	1	1	1	84.800	4,9	Rendah
19.	1	3	2	209.000	4,7	Rendah
20.	1	1	1	89.000	4,9	Tinggi
21.	2	3	0	159.000	5	Rendah
22.	2	3	0	159.000	4,9	Rendah
23.	2	1	1	89.000	4,9	Tinggi
24.	2	3	0	179.000	4,8	Rendah
25.	2	1	1	89.000	4,6	Tinggi
26.	2	1	1	69.000	4,8	Tinggi
27.	2	1	1	89.000	4,9	Tinggi
28.	2	0	0	79.000	4,8	Rendah
29.	3	1	1	130.000	4,8	Rendah



30.	3	1	0	45.000	4,8	Tinggi
31.	3	3	0	89.000	4,9	Tinggi
32.	3	0	0	55.000	4,9	Tinggi
33.	3	3	0	80.000	4,9	Tinggi
34.	3	1	1	25.900	4,8	Tinggi
35.	3	2	1	129.000	4,8	Rendah
36.	3	2	0	84.000	5	Rendah
37.	3	0	0	69.000	4,9	Tinggi
38.	3	0	3	89.000	4,8	Tinggi
39.	3	3	3	189.000	4,9	Tinggi
40.	4	0	3	89.000	4,8	Tinggi
41.	4	2	2	79.000	4,9	Rendah
42.	4	3	3	79.000	4,9	Rendah
43.	4	2	1	89.000	4,6	Rendah
44.	4	2	1	89.000	4,7	Rendah
45.	4	0	3	35.000	4,8	Tinggi
46.	5	2	2	40.000	4,9	Tinggi
47.	5	0	3	38.710	4,8	Tinggi
48.	5	2	2	35.000	4,8	Tinggi
49.	5	3	3	35.000	4,7	Tinggi
50.	5	2	2	35.000	5	Tinggi
51.	5	2	3	30.000	5	Rendah
52.	5	1	3	30.000	5	Rendah
53.	5	1	1	55.000	5	Rendah
54.	5	0	3	30.000	5	Rendah
55.	6	1	1	325.000	4,9	Tinggi
56.	6	1	1	150.000	5	Rendah
57.	6	1	1	300.000	5	Rendah
58.	6	1	1	300.000	5	Rendah
59.	6	1	1	275.000	5	Rendah
60.	6	1	1	100.000	4,9	Rendah

61.	6	1	1	325.000	4,9	Rendah
62.	6	1	1	300.000	4,9	Rendah
63.	7	2	1	199.000	5	Rendah
64.	7	2	0	199.000	4,9	Rendah
65.	7	2	2	199.000	5	Renda
66.	7	2	1	199.000	5	Rendah
67.	7	1	1	229.000	4,9	Rendah
68.	7	1	1	349.000	4,8	Rendah
69.	8	1	1	25.000	4,9	Tinggi
70.	8	2	1	35.000	4,9	Rendah
71.	8	2	1	25.000	4,9	Rendah
72.	8	1	1	45.000	5	Rendah
73.	8	1	0	40.000	4,9	Rendah
74.	8	2	0	35.000	4,8	Rendah
75.	8	1	2	45.000	4,9	Rendah
76.	9	0	3	39.900	4,8	Tinggi
77.	9	2	1	44.990	4,9	Rendah
78.	9	2	1	52.990	4,8	Tinggi
79.	9	0	3	49.500	4,8	Tinggi
80.	9	1	1	35.000	4,8	Rendah
81.	9	0	0	37.500	4,7	Tinggi
82.	9	3	3	65.000	4,7	Rendah
83.	10		2	210.000	4,9	Rendah
84.	10	3	2	270.000	5	Rendah
85.	10	3	3	190.000	4,9	Rendah
86.	10	2	2	170.000	4,8	Rendah
87.	11	0	0	105.000	4,9	Rendah
88.	11	0	0	105.000	4,9	Rendah
89.	11	2	2	110.000	4,9	Rendah
90.	11	2	2	85.000	4,9	Rendah
91.	11	1	1	129.900	4,9	Rendah

92.	11	0	3	110.000	5	Rendah
93.	11	0	0	99.000	5	Rendah
94.	12	1	1	215.000	4,9	Rendah
95.	12	2	2	195.000	5	Rendah
96.	12	1	1	325.000	5	Rendah
97.	12	1	1	325.000	5	Rendah
98.	12	1	1	325.000	4,9	Rendah
99.	12	1	1	299.000	4,9	Rendah
100.	12	1	1	195.000	5	Rendah

## Lampiran 6. Rules Hasil *Decision Tree*

Rules untuk Hasil Kelas = Tinggi
1. IF nilai Harga > 91195 and nilai Jenis > 2.5 and nilai Brand ≤ 2.5 THEN Kelas Tinggi.
2. IF nilai Harga > 91195 and nilai Jenis > 2.5 and nilai Brand > 2.5 and nilai Bintang > 4.95 THEN Kelas Tinggi.
3. IF nilai Harga > 91195 and nilai Jenis ≤ 2.5 and nilai Harga ≤ 312500 THEN Kelas Tinggi.
4. IF nilai Harga > 91195 and nilai Jenis ≤ 2.5 and nilai Harga > 312500 and nilai Brand > 6.5 THEN Kelas Tinggi.
5. IF nilai Harga > 91195 and nilai Jenis ≤ 2.5 and nilai Harga > 312500 and nilai Brand ≤ 6.5 THEN Kelas Tinggi.
6. IF nilai Harga ≤ 91195 and nilai Bintang > 4.95 and nilai Jenis ≤ 1.5 THEN Kelas Tinggi.
7. IF nilai Harga ≤ 91195 and nilai Bintang > 4.95 and nilai Jenis > 1.5 and nilai Brand ≤ 4 THEN Kelas Tinggi.
8. IF nilai Harga ≤ 91195 and nilai Bintang > 4.95 and nilai Jenis > 1.5 and nilai Brand > 4 and nilai Harga ≤ 32500 THEN Kelas Tinggi.
9. IF nilai Harga ≤ 91195 and nilai Bintang ≤ 4.95 and nilai Jenis > 1.5 and nilai Brand > 3.5 and nilai Harga > 58950 THEN Kelas Tinggi.
10. IF nilai Harga ≤ 91195 and nilai Bintang ≤ 4.95 and nilai Jenis > 1.5 and nilai Brand > 3.5 and nilai Harga ≤ 58950 and nilai Bahan ≤ 1.5 and nilai Harga ≤ 48945 THEN Kelas Tinggi.
11. IF nilai Harga ≤ 91195 and nilai Bintang ≤ 4.95 and nilai Jenis ≤ 1.5 and nilai Brand > 1.5 and nilai Brand > 6.5 and nilai Jenis > 0.5 and nilai Harga > 30000 THEN Kelas Tinggi.
12. IF nilai Harga ≤ 91195 and nilai Bintang ≤ 4.95 and nilai Jenis ≤ 1.5 and nilai Brand > 1.5 and nilai Brand ≤ 6.5 and nilai Bahan ≤ 0.5 and nilai Brand ≤ 2.5 THEN Kelas Tinggi.
13. IF nilai Harga ≤ 91195 and nilai Bintang ≤ 4.95 and nilai Jenis ≤ 1.5 and nilai Brand ≤ 1.5 and nilai Bintang ≤ 4.8 THEN Kelas Tinggi.

14. IF nilai Harga  $\leq$  91195 and nilai Bintang  $\leq$  4.95 and nilai Jenis  $\leq$  1.5 and nilai Brand  $\leq$  1.5 and nilai Bintang  $>$  4.8 and nilai Harga  $>$  76315 and nilai Harga  $\leq$  86900 THEN Kelas Tinggi.

**Rules untuk Hasil Kelas = Rendah**

1. IF nilai Harga  $>$  91195 and nilai Jenis  $>$  2.5 and nilai Brand  $>$  2.5 and nilai Bintang  $\leq$  4.95 THEN Kelas Rendah.
2. IF nilai Harga  $\leq$  91195 and nilai Bintang  $>$  4.95 and nilai Jenis  $>$  1.5 and nilai Brand  $>$  4 and nilai Harga  $>$  32500 THEN Kelas Rendah.
3. IF nilai Harga  $\leq$  91195 and nilai Bintang  $\leq$  4.95 and nilai Jenis  $>$  1.5 and nilai Brand  $\leq$  3.5 THEN Kelas Rendah.
4. IF nilai Harga  $\leq$  91195 and nilai Bintang  $\leq$  4.95 and nilai Jenis  $>$  1.5 and nilai Brand  $>$  3.5 and nilai Harga  $\leq$  58950 and nilai Bahan  $>$  1.5 THEN Kelas Rendah.
5. IF nilai Harga  $\leq$  91195 and nilai Bintang  $\leq$  4.95 and nilai Jenis  $>$  1.5 and nilai Brand  $>$  3.5 and nilai Harga  $\leq$  58950 and nilai Bahan  $\leq$  1.5 and nilai Harga  $>$  48945 THEN Kelas Rendah.
6. IF nilai Harga  $\leq$  91195 and nilai Bintang  $\leq$  4.95 and nilai Jenis  $\leq$  1.5 and nilai Brand  $>$  1.5 and nilai Brand  $>$  6.5 and nilai Jenis  $\leq$  0.5 THEN Kelas Rendah.
7. IF nilai Harga  $\leq$  91195 and nilai Bintang  $\leq$  4.95 and nilai Jenis  $\leq$  1.5 and nilai Brand  $>$  1.5 and nilai Brand  $>$  6.5 and nilai Jenis  $>$  0.5 and nilai Harga  $\leq$  30000 THEN Kelas Rendah.
8. IF nilai Harga  $\leq$  91195 and nilai Bintang  $\leq$  4.95 and nilai Jenis  $\leq$  1.5 and nilai Brand  $>$  1.5 and nilai Brand  $\leq$  6.5 and nilai Bahan  $>$  0.5 THEN Kelas Rendah.
9. IF nilai Harga  $\leq$  91195 and nilai Bintang  $\leq$  4.95 and nilai Jenis  $\leq$  1.5 and nilai Brand  $>$  1.5 and nilai Brand  $\leq$  6.5 and nilai Bahan  $\leq$  0.5 and nilai Brand  $>$  2.5 THEN Kelas Rendah.

10. IF nilai Harga  $\leq$  91195 and nilai Bintang  $\leq$  4.95 and nilai Jenis  $\leq$  1.5 and nilai Brand  $\leq$  1.5 and nilai Bintang  $>$  4.8 and nilai Harga  $\leq$  76315 THEN Kelas Rendah.
11. IF nilai Harga  $\leq$  91195 and nilai Bintang  $\leq$  4.95 and nilai Jenis  $\leq$  1.5 and nilai Brand  $\leq$  1.5 and nilai Bintang  $>$  4.8 and nilai Harga  $>$  76315 and nilai Harga  $>$  86900 THEN Kelas Rendah.

## RIWAYAT HIDUP



Faridatul Husna, lahir di Jepara pada tanggal 12 Februari 2000. Putri pertama dari (Alm) Bapak Suyanto dan Ibu Siti Asiyah. dilahirkan dan dibesarkan di rumah sederhana terletak di Dusun Metawar, Desa Ujungwatu, Kecamatan Donorojo, Kabupaten Jepara.

Perempuan dengan sapaan akrab Husna, telah menempuh pendidikan formal mulai dari RA Al Ikhlas pada tahun 2005, dilanjutkan dengan pendidikan dasar di MI Nurul Huda dan lulus pada tahun 2011, selanjutnya menempuh jenjang menengah pertama di SMPN 1 Donorojo dan lulus pada tahun 2014, kemudian menempuh jenjang menengah atas di SMAN 1 Tayu dan lulus pada tahun 2017. Pada tahun yang sama, ia menempuh kuliah di Universitas Islam Negeri Maulana Malik Ibrahim pada Program Studi Matematika.

Selama menjadi mahasiswa, penulis berperan aktif dalam mengembangkan kemampuan akademiknya dengan menjadi asisten laboratorium selama 1 semester. Selain itu, penulis juga berperan aktif dalam kepengurusan HMJ (Himpunan Mahasiswa Jurusan) Integral Matematika selama 2 periode.



KEMENTERIAN RI  
UNIVERSITAS ISLAM NEGERI  
MAULANA MALIK IBRAHIM MALANG  
FAKULTAS SAINS DAN TEKNOLOGI  
Jl. Gajayana No. 50 Dinoyo Malang Telp/Fax.(0341)558933

### BUKTI KONSULTASI SKRIPSI

Nama : Faridatul Husna  
NIM : 17610069  
Fakultas/ Program Studi : Sains dan Teknologi/ Matematika  
Judul Skripsi : Implementasi Data Mining Menggunakan Algoritma  
C4.5 pada Klasifikasi Penjualan Hijab  
Pembimbing I : Dr. Hairur Rahman, M.Si  
Pembimbing II : Juhari, M.Si

No.	Tanggal	Hal	Tanda Tangan	
1.	10 Maret 2021	Konsultasi Bab I dan Bab II	1. <i>ff</i>	
2.	25 Maret 2021	Konsultasi Agama Bab I		2. <i>ff</i>
3.	03 April 2021	Revisi Bab I dan Bab II	3. <i>ff</i>	
4.	07 Mei 2021	ACC Bab I dan Bab II		4. <i>ff</i>
5.	30 Mei 2021	Konsultasi Agama Bab II	5. <i>ff</i>	
6.	31 Mei 2021	Konsultasi Bab III		6. <i>ff</i>
7.	14 Juni 2021	Revisi Bab III	7. <i>ff</i>	
8.	15 Juni 2021	ACC Bab III		8. <i>ff</i>
9.	17 November 2021	Konsultasi Pemrograman	9. <i>ff</i>	
10.	26 November 2021	Konsultasi Bab IV		10. <i>ff</i>
11.	27 November 2021	Konsultasi Integrasi Keislaman	11. <i>ff</i>	
12.	29 November 2021	ACC Bab IV		12. <i>ff</i>
13.	30 November 2021	ACC Agama Keseluruhan	13. <i>ff</i>	

Malang, 28 Desember 2021

Mengetahui,  
Ketua Program Studi Matematika

*Elly Susanti*  
Dr. Elly Susanti, M.Sc

NIP. 19741129 200012 2 005