Falerina, Gabriel, Dwi, Melki, Heru

GALLEY JUISI 1386 - Falentina et al

Document Details

Submission ID
trn:oid:::1:3442651177

Submission Date
Dec 12, 2025, 3:10 PM GMT+7

Download Date
Dec 18, 2025, 10:21 PM GMT+7

File Name
GALLEY_JUISI_1386_-_Falentina_et_al_n.p.docx

File Size
9.6 MB

15 Pages

5,955 Words

39,150 Characters

14% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Quoted Text

Top Sources

12% Internet sources
8% Publications
4% Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Top Sources

12% Internet sources
8% Publications
4% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Internet	
ejurnal.provisi.ac.id		2%
2	Student papers	
Universitas PGRI Semarang		2%
3	Internet	
docplayer.info		<1%
4	Internet	
kc.umn.ac.id		<1%
5	Internet	
repository.usd.ac.id		<1%
6	Student papers	
UM Surabaya		<1%
7	Publication	
Luhur Pambudi, Sefrika Sefrika. "Klasifikasi Dampak dan Kondisi Pasien Hepatitis ...		<1%
8	Student papers	
Universitas Gunadarma		<1%
9	Publication	
Patmawati Hasan, Heru Sutejo. "SISTEM PENDUKUNG KEPUTUSAN SELEKSI PELA...		<1%
10	Internet	
www.researchgate.net		<1%
11	Internet	
menukuliner.net		<1%

12	Student papers	
Universitas Komputer Indonesia		<1%
13	Publication	
Sofia Dewi, Oktaviawati Oktaviawati. "Penerapan Algoritma C4.5 untuk Pehaman...		<1%
14	Student papers	
Universitas Dian Nuswantoro		<1%
15	Student papers	
Loughborough University		<1%
16	Internet	
ojs.amikomsolo.ac.id		<1%
17	Internet	
dabar.srce.hr		<1%
18	Internet	
es.scribd.com		<1%
19	Internet	
jurnal.fikom.umi.ac.id		<1%
20	Internet	
www.grafiati.com		<1%
21	Internet	
journal.nurulfikri.ac.id		<1%
22	Internet	
journal.ummat.ac.id		<1%
23	Internet	
www.ioinformatic.org		<1%
24	Internet	
zenodo.org		<1%
25	Publication	
Elsa Paskalis Krisda Orpa, Eva Faja Ripanti, Tursina Tursina. "Model Prediksi Awal ...		<1%

26	Internet	
core.ac.uk		<1%
27	Internet	
ejournal.upbatam.ac.id		<1%
28	Internet	
ejurnal.unikarta.ac.id		<1%
29	Internet	
journal.peradaban.ac.id		<1%
30	Internet	
jsisfotek.org		<1%
31	Internet	
u-read.unsri.ac.id		<1%
32	Publication	
Christ Mario, Ryan Randy Suryono. "Public Sentiment Analysis on Dirty Vote Movi...		<1%
33	Publication	
Meilin Widyastuti, Agnes Gracella Fepdiani Simanjuntak, Dedy Hartama, Agus Per...		<1%
34	Publication	
Rosyid Eko Nugroho, Wisnu Yogi Pamungkas, Jajam Haerul Jaman. "PENDETEKSI ...		<1%
35	Internet	
ejournal.itn.ac.id		<1%
36	Internet	
journal.thamrin.ac.id		<1%
37	Internet	
jurnal-tmit.com		<1%
38	Internet	
jurnal.murnisadar.ac.id		<1%
39	Internet	
jurnal.ulb.ac.id		<1%

40 Internet

tabloidjubi.com <1%

41 Internet

tunasbangsa.ac.id <1%

42 Publication

Aji Primajaya, Betha Nurina Sari, Ahmad Khusaeri. "Prediksi Potensi Kebakaran H... <1%

43 Publication

Muhammad Rizki Fahdia, Dwiza Riana, Fachri Amsury, Irwansyah Saputra, Nanan... <1%

44 Publication

Dadang Iskandar Mulyana, Muhammad Hafiz Siregar. "Deteksi Dini Siswa Korban... <1%

45 Publication

Fery bayu aji, Fajri Rakhmat Umbara, Fatan Kasyidi. "KLASIFIKASI RISIKO KEMAT... <1%

turnitin Page 7 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

OPEN ACCESS

JURNAL ILMIAH SISTEM INFORMASI (JUSI)
VOL 4, No. 2, Mei 2025, pp. XX - XX
p-ISSN : 2803-1507 (media cetak) | e-ISSN : 2803-1531 (media online)
Link Publikasi: https://ejurnal.provisi.ac.id/index.php/JUISI/issue/view/134

Analysis of Data Mining Implementation for Classifying Best-Selling Food Sale Using the Decision Tree (C4.5) Algorithm

Falerina Gita Falentina*[1], Gabriel Yohan Yoseph Wabdaron[2], Dwi Andiyani[3], Melki Sendoni Wondiwoi[4], Heru Sutejo[5]
[1,2,3,4,5] Universitas Sepuluh Nopember Papua
Jl. Ardipura Raya No.22B, 99222, Ardipura, Distrik Jayapura Selatan, Kota Jayapura, Papua, Indonesia
e-mail: falerinagita28@gmail.com*[1], wapdarongabriel@gmail.com[2], dwiandianiii03@gmail.com[3], Melkyshow@gmail88269.com[4], heru.sutejo01@gmail.com[5]

ARTICLE INFO

ABSTRACT

The culinary business sector continues to grow rapidly, creating a strong need for data-driven decision-making to support operational efficiency. This study aims to classify the sales performance of menu items at Warung Makan Lalapan Haris by applying the C4.5 decision tree algorithm and the KDD methodology. A total of 500 data records containing attributes such as menu type, number of customers, number of items sold, and sales status were processed through several stages, including data selection, preprocessing, transformation, data mining, and evaluation. The decision tree model was built using RapidMiner 2026.0.1 with a data split of 70% for training and 30% for testing. The results indicate that the C4.5 algorithm successfully formed a classification structure that categorizes menu items into Best Seller, Medium Seller, and Low Seller groups. Fried Chicken was consistently identified as the Best Seller, Grilled Chicken as the Medium Seller, while Fried Catfish and Grilled Catfish were classified as Low Seller items with more complex splitting patterns primarily influenced by the number of customers. The evaluation results show an accuracy of 84%, with perfect precision and recall for Fried Chicken and Grilled Chicken, while the performance for Fried Catfish and Grilled Catfish varied. These findings demonstrate that the C4.5 decision tree algorithm is effective for analyzing sales patterns and can assist business owners in planning inventory and optimizing menu management strategies.

Keywords: *C4.5 Decision Tree Algorithm, Data Mining, Sales Classification, RapidMiner.*

Abstrak

Sektor bisnis kuliner terus berkembang pesat, menciptakan kebutuhan yang kuat akan pengambilan keputusan berbasis data untuk mendukung efisiensi operasional. Penelitian ini bertujuan untuk mengklasifikasikan kinerja penjualan item menu di Warung Makan Lalapan Haris dengan menerapkan algoritma pohon keputusan C4.5 dan metodologi KDD. Sebanyak 500 data record yang berisi atribut seperti jenis menu, jumlah pelanggan, jumlah barang terjual, dan status penjualan diproses melalui beberapa tahap, meliputi pemilihan data, praproses, transformasi, penggalian data, dan evaluasi. Model pohon keputusan

DOI: https://doi.org/10.51903/f2jegk76

Jurnal Ilmiah Sistem Informasi

turnitin Page 7 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

turnitin Page 8 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

p-ISSN : 2803-1507 e-ISSN : 2803-1531

dibangun menggunakan RapidMiner 2026.0.1 dengan pembagian data 70% untuk pelatihan dan 30% untuk pengujian. Hasil penelitian menunjukkan bahwa algoritma C4.5 berhasil membentuk struktur klasifikasi yang mengkategorikan item menu ke dalam kelompok *Best Seller, Medium Seller, dan Low Seller*. Ayam Goreng secara konsisten diidentifikasi sebagai *Best Seller*, Ayam Bakar sebagai *Medium Seller*, sementara Lele Goreng dan Lele Bakar diklasifikasikan sebagai *item Low Seller* dengan pola pembagian yang lebih kompleks yang terutama dipengaruhi oleh jumlah pelanggan. Hasil evaluasi menunjukkan akurasi 84%, dengan presisi dan recall sempurna untuk Ayam Goreng dan Ayam Bakar, sementara performa untuk Lele Goreng dan Lele Bakar bervariasi. Temuan ini menunjukkan bahwa algoritma pohon keputusan C4.5 efektif untuk menganalisis pola penjualan dan dapat membantu pemilik bisnis dalam merencanakan inventaris dan mengoptimalkan strategi manajemen menu.

Kata Kunci: *Algoritma Decision Tree C4.5, Data Mining, Klasifikasi Penjualan, RapidMiner.*

1. PENDAHULUAN

Usaha kuliner merupakan salah satu sektor yang memiliki peran penting dalam perekonomian Indonesia karena mampu berkembang secara konsisten dan tetap bertahan meskipun menghadapi berbagai tantangan. Industri ini diminati oleh berbagai pelaku usaha, mulai dari restoran besar hingga warung makan tradisional, seiring meningkatnya kebutuhan masyarakat terhadap makanan siap saji yang praktis, terjangkau, dan sesuai dengan selera lokal [1]. Kehadiran usaha kuliner juga mencerminkan kebutuhan konsumen terhadap variasi makanan yang praktis, terjangkau, dan sesuai dengan selera lokal. Dengan semakin tingginya persaingan, pelaku usaha perlu memiliki strategi pengelolaan yang lebih efektif agar dapat mempertahankan keberlanjutan usahanya [2].

Warung Makan Lalapan Haris merupakan salah satu usaha kuliner yang menawarkan hidangan dengan cita rasa khas lokal, seperti ayam goreng, ayam bakar, lele goreng, dan lele bakar. Meskipun menu yang disediakan relatif sederhana, permintaan pelanggan terhadap setiap menu tidak selalu stabil. Fluktuasi permintaan ini menyebabkan pemilik usaha kerap mengalami kesulitan dalam merencanakan stok bahan baku maupun mengatur operasional harian, terutama saat terjadi lonjakan jumlah pelanggan [3]. Variasi jumlah pelanggan pada hari-hari tertentu juga dapat memengaruhi perputaran stok sehingga diperlukan pemantauan yang lebih cermat terhadap pola kunjungan.

Dalam era digital saat ini, banyak pelaku usaha mulai memanfaatkan data sebagai dasar pengambilan keputusan untuk meningkatkan efisiensi operasional [4]. Data penjualan yang dimiliki Warung Makan Lalapan Haris sebenarnya dapat dianalisis lebih jauh untuk memahami pola pembelian dan preferensi pelanggan. Berbagai penelitian menunjukkan bahwa teknik data mining dapat membantu mendeteksi kecenderungan pelanggan secara lebih akurat, khususnya melalui metode klasifikasi, Salah satunya adalah penelitian oleh Ahmad & Tresna dalam klasifikasi Data Kandidat Tenaga Kerja [5]. Salah satu metode yang banyak digunakan karena mudah dipahami dan menghasilkan aturan yang jelas adalah algoritma decision tree C4.5, yang mampu memetakan hubungan antara atribut penentu dan kategori hasil dalam bentuk struktur pohon keputusan.

Penelitian ini bertujuan menerapkan algoritma decision tree C4.5 untuk menganalisis data penjualan Warung Makan Lalapan Haris sehingga dapat mengidentifikasi menu yang berpotensi menjadi *best seller*. Model klasifikasi yang dihasilkan diharapkan mampu menjadi dasar pengambilan keputusan yang lebih terukur dalam perencanaan stok bahan baku serta pengelolaan operasional harian [6]. Dari sisi akademik, penelitian ini berkontribusi pada pengembangan literatur mengenai penerapan metode klasifikasi dalam konteks usaha kuliner tradisional [7]. Sementara itu, kontribusi praktis penelitian ini terletak pada penyediaan solusi berbasis data yang dapat mendukung peningkatan kinerja penjualan serta efisiensi manajemen operasional pada Warung Makan Lalapan Haris.

2. TINJAUAN PUSTAKA

2.1. *Data Mining*

Data mining banyak dibahas dalam literatur sebagai pendekatan analitik untuk mengolah dan mengekstraksi pola dari data berskala besar [8], [9], [10]. Menurut [11], pendekatan ini mencakup tahapan seperti pembersihan, integrasi, seleksi, transformasi data, serta analisis dan interpretasi hasil. Dalam konteks bisnis, literatur tersebut menyebutkan bahwa *data mining* digunakan untuk mendukung pengambilan keputusan berbasis data. Pada bidang kuliner, *data mining* dijelaskan sebagai teknik yang membantu mengidentifikasi

turnitin Page 8 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

turnitin Page 9 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

p-ISSN : 2803-1507 e-ISSN : 2803-1531

kecenderungan penjualan sehingga menu yang berpotensi memiliki permintaan tinggi dapat diprediksi dan proses operasional dapat ditingkatkan.

2.2. *Knowledge Discovery in Database (KDD)*

KDD diperkenalkan dalam berbagai studi sebagai kerangka kerja sistematis untuk menghasilkan informasi dari data mentah melalui tahapan yang terstruktur. Menurut [12], kerangka ini terdiri dari proses pemilihan, pembersihan, transformasi data, penerapan algoritma, dan interpretasi pola. Studi yang dibahas pada [13] menjelaskan bahwa tiap tahapan memiliki peran dalam menjaga kualitas informasi yang diekstraksi. Literatur yang sama menggambarkan *KDD* sebagai pendekatan yang memungkinkan analisis dilakukan secara bertahap sehingga kualitas hasil tetap terkontrol.

2.3. *Rapid Miner*

RapidMiner sering digunakan dalam penelitian *data mining* karena menyediakan fasilitas pemodelan yang mudah dioperasikan melalui antarmuka visual. Menurut [14], perangkat ini mendukung proses pra-pemrosesan data, pembangunan model, dan evaluasi performa menggunakan berbagai metrik seperti akurasi, *precision*, *recall*, dan *confusion matrix*. Studi tersebut juga menyebutkan bahwa versi 2026.0.1 menawarkan peningkatan performa komputasi, stabilitas antarmuka, serta penambahan operator yang kompatibel dengan algoritma modern, sehingga proses analisis dapat dilakukan secara lebih efisien. Penggunaan *RapidMiner* dalam berbagai penelitian juga menunjukkan bahwa perangkat ini mampu mendukung eksperimen model secara berulang melalui pengaturan alur proses yang terstruktur.

2.4. *Split Data*

Teknik *split data* diperkenalkan dalam literatur sebagai metode penting untuk memisahkan fase pelatihan model dengan fase pengujiannya. Menurut [15], pembagian *dataset* menjadi data pelatihan dan pengujian digunakan untuk menilai kemampuan model dalam melakukan generalisasi terhadap data baru. Studi tersebut menjelaskan bahwa penentuan proporsi pembagian dilakukan agar distribusi karakteristik data pada kedua subset tetap seimbang, sehingga evaluasi performa model menjadi lebih objektif dan representatif. Beberapa penelitian juga menggambarkan bahwa teknik ini membantu peneliti memahami perubahan performa model ketika diuji pada variasi data yang berbeda.

2.5. *Algoritma Decision Tree C4.5*

Algoritma C4.5 menjadi salah satu metode klasifikasi yang banyak dibahas dalam penelitian karena menghasilkan pohon keputusan yang mudah dipahami. Menurut [16], algoritma ini mampu mengolah data kategorikal maupun numerik serta menangani nilai hilang. Studi pada [17] menunjukkan bahwa proses pemangkasan (*pruning*) digunakan untuk mengurangi risiko *overfitting*. Dalam konteks analisis penjualan, [18] menggambarkan penerapan C4.5 untuk menelusuri hubungan antaratribut sehingga faktor yang berpengaruh terhadap permintaan dapat dikenali. Struktur pohon dibangun melalui perhitungan *gain ratio*, di mana atribut dengan nilai tertinggi ditetapkan sebagai simpul utama sebelum pemilihan atribut berikutnya dilakukan.

3. METODOLOGI PENELITIAN

3.1. *Metode Penelitian*

Penelitian ini menggunakan metode KDD *(Knowledge Discovery in Database),* yaitu proses penemuan pola atau pengetahuan baru dari data melalui beberapa tahapan sistematis. Metode ini dipilih karena sesuai untuk penelitian yang bertujuan menganalisis data penjualan dan membangun model prediksi menggunakan algoritma data mining [19]. Metode ini memungkinkan proses analisis dilakukan secara terstruktur sehingga setiap tahapan dapat diidentifikasi secara jelas. Pendekatan tersebut membantu memastikan bahwa data yang dianalisis telah melalui proses persiapan yang memadai untuk mendukung kualitas hasil pemodelan.

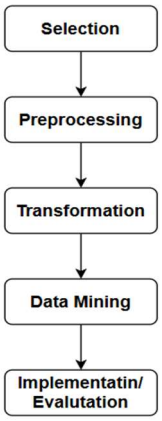
3.2. *Tahapan KDD*

Proses KDD dalam penelitian ini terdiri dari lima tahap, yaitu *Selection*, *Preprocessing*, *Transformation*, Data Mining, dan *Interpretation/Evaluation*. Setiap tahap memiliki fungsi spesifik yang saling berkaitan dalam menghasilkan pola informasi yang akurat dari data yang dianalisis. Struktur tahapan tersebut membantu memandu proses penelitian agar berjalan secara sistematis dan sesuai prinsip pengolahan data yang benar. Tahapan KDD dapat dilihat pada Gambar 1.

turnitin Page 9 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

turnitin Page 10 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

p-ISSN : 2803-1507 e-ISSN : 2803-1531



Gambar 1. Tahapan Metodologi Knowledge Discovery in Database

3.2.1 *Selection (Seleksi Data)*

Pada tahap ini dilakukan pemilihan data yang relevan dari keseluruhan data penjualan [20]. Data yang digunakan dalam penelitian ini berjumlah 500 data *record*, yang berisi informasi mengenai: Menu, Jumlah Pelanggan, Jumlah Terjual, Status Menu (*Low*/*Medium*/*Best Seller*). Atribut-atribut tersebut dipilih karena memiliki pengaruh langsung terhadap proses klasifikasi. Pemilihan atribut yang tepat membantu memastikan bahwa proses klasifikasi dapat berjalan lebih optimal berdasarkan informasi yang tersedia.

3.2.2 *Preprocessing (Prapemrosesan Data)*

Pada tahap *pre-processing* dilakukan proses pembersihan data untuk memastikan bahwa dataset berada dalam kondisi optimal sebelum masuk ke tahap pemodelan. Proses ini mencakup identifikasi dan penanganan *missing value*, memperbaiki ketidakkonsistenan antar atribut, serta menstandarkan format data [21]. proses pembersihan data dilakukan bersamaan dengan mekanisme *split data* menggunakan metode *shuffled sampling*. Melalui tahapan ini, data yang digunakan menjadi lebih bersih, rapi, dan siap untuk diolah pada proses analisis dan pembangunan model *Decision Tree*. Dengan kualitas data yang baik, akurasi model dapat ditingkatkan dan hasil prediksi menjadi lebih dapat diandalkan .

3.2.3 *Transformation (Transformasi Data)*

Data yang telah dibersihkan kemudian diubah ke dalam bentuk yang sesuai untuk Algoritma *Decision Tree* C4.5. Proses yang dilakukan, Mengonversi atribut Menu menjadi *polynominal*, Mengelompokkan variabel numerik, Mengubah struktur dataset agar kompatibel dengan operator RapidMiner, Transformasi ini memastikan setiap atribut dapat diproses oleh model klasifikasi [22]. Proses transformasi dilakukan agar struktur data berada dalam format yang sesuai dengan kebutuhan algoritma sehingga langkah pemodelan dapat berjalan lebih efisien. Data Penjualan Dapat dilihat pada Gambar 2.

	Menu *polynominal*	Jumlah_Pelanggan *integer*	Jumlah_Terjual *integer*	Status_Menu *polynominal*
1	Lele Goreng	15	15	Low Seller
2	Lele Goreng	13	18	Low Seller
3	Lele Goreng	18	27	Low Seller
4	Ayam Bakar	18	40	Medium Seller
5	Lele Bakar	20	20	Low Seller
6	Ayam Bakar	15	28	Medium Seller
7	Ayam Bakar	17	36	Medium Seller
8	Lele Goreng	9	14	Low Seller
9	Lele Goreng	7	8	Low Seller
10	Ayam Goreng	15	51	Best Seller
11	Lele Goreng	16	23	Low Seller
12	Lele Bakar	17	25	Low Seller

Gambar 2. Data Penjualan Menu

turnitin Page 10 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

turnitin Page 11 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

p-ISSN : 2803-1507 e-ISSN : 2803-1531

3.2.4 *Data Mining (Penerapan Algoritma Decision Tree C4.5)*

Pada bagian ini dilakukan penerapan metode data mining menggunakan algoritma *Decision Tree* C4.5. Algoritma ini membangun model klasifikasi melalui proses perhitungan, *Entropy, Gain, Split Information, Gain Ratio* [23]. Proses pemodelan dilakukan dengan bantuan perangkat lunak RapidMiner versi 2026.0.1, di mana dataset dibagi menggunakan metode *Split Data*. Pembagian dilakukan sebesar, 70% sebagai data training untuk membangun model, dan 30% sebagai data testing untuk mengukur performa model. Hasil proses data mining menghasilkan model *Decision Tree* yang mampu mengklasifikasikan menu ke dalam kategori *Low Seller, Medium Seller,* atau *Best Seller* berdasarkan pola historis penjualan.

Pada tahap pemilihan atribut terbaik dalam Algoritma *Decision Tree* C4.5, dilakukan perhitungan nilai *entropy, gain, split information, dan gain ratio*. Perhitungan ini bertujuan untuk menentukan atribut yang memiliki kemampuan klasifikasi paling tinggi terhadap variabel keputusan (*class label*). Semakin besar nilai *gain ratio* yang dihasilkan oleh suatu atribut, maka semakin layak atribut tersebut digunakan sebagai pemisah (*node*) pada *decision tree*. Dataset yang digunakan terdiri dari lima sampel menu beserta kategori hasil penjualannya. Data tersebut disajikan secara lengkap pada Tabel 1.

Tabel 1. Lima Sampel Menu

Menu	Jumlah Pelanggan	Jumlah Terjual	Status
Ayam Bakar	8	17	Medium Seller
Lele Bakar	5	6	Low Seller
Ayam Goreng	19	49	Best Seller
Lele Goreng	10	15	Low seller
Lele Goreng	9	13	Low Seller

Jumlah kelas pada dataset yaitu *Best Seller* sebanyak 1 data, *Medium Seller* sebanyak 1 data, dan *Low Seller* sebanyak 3 data. Berdasarkan distribusi kelas tersebut, nilai *entropy* total dataset dihitung dengan menggunakan rumus (1) untuk mengetahui tingkat ketidakpastian awal sebelum proses pemisahan atribut dilakukan. Perhitungan ini berfungsi sebagai dasar dalam menentukan atribut mana yang paling sesuai untuk dijadikan kandidat node awal dalam pohon keputusan. Hasil lengkap perhitungan tersebut ditampilkan untuk kejelasan pada Gambar 3.

Rumus *Entropy*:

$$Entropy\ (S) = -\sum_{i=1}^{n} P_i \ \times \log_2 P_i \qquad (1)$$

Perhitungan :
$p_{Low} = 3/5 = 0.6$
$p_{Medium} = 1/5 = 0.2$
$p_{Best} = 1/5 = 0.2$

$$Entropy(S) = -[0.6\log_2 0.6 + 0.2\log_2 0.2 + 0.2\log_2 0.2]$$

Hitung log:
$\log_2 0.6 = -0.736$
$\log_2 0.2 = -2.322$

$$Entropy(S) = -[(0.6)(-0.736) + (0.2)(-2.322) + (0.2)(-2.322)]$$
$$Entropy(S) = -[-0.4416 - 0.4644 - 0.4644]$$
$$Entropy(S) = 1.3704$$

Kelas	Jumlah	Probabilitas	-p log2 p
Low Seller	3	0.6	0.4416
Medium Sell	1	0.2	0.4644
Best Seller	1	0.2	0.4644
Total	5		1.3704

Gambar 3. Hasil Tabel Perhitungan Entropy

turnitin Page 11 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

turnitin Page 12 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

p-ISSN : 2803-1507 e-ISSN : 2803-1531

Selanjutnya dilakukan perhitungan *entropy* untuk setiap nilai pada atribut Menu. Tahap ini dilakukan untuk mengevaluasi apakah setiap kategori menu memiliki variasi kelas atau hanya satu jenis status penjualan pada kelompok datanya. Perhitungan ini penting karena *entropy* nol menunjukkan bahwa atribut tersebut sepenuhnya homogen sehingga tidak memiliki ketidakpastian. Rumus dan hasil perhitungan ditunjukkan pada persamaan (2).

$$Entropy(Menu) = \frac{1}{5}(0) + \frac{1}{5}(0) + \frac{1}{5}(0) + \frac{2}{5}(0) \quad (2)$$

Ayam Bakar (1 data → *Medium Seller*) *Entropy* = 0
Lele Bakar (1 data → *Low Seller*) *Entropy* = 0
Ayam Goreng (1 data → *Best Seller*) *Entropy* = 0
Lele Goreng (2 data → *Low Seller, Low Seller*) *Entropy* = 0

Seluruh kategori menu (Ayam Bakar, Lele Bakar, Ayam Goreng, dan Lele Goreng) hanya memiliki satu kelas dominan untuk masing-masing kelompok datanya, sehingga seluruh nilai *entropy* kategori atribut Menu bernilai 0. Hal ini menunjukkan bahwa tidak ada ketidakpastian pada pembagian berdasarkan atribut Menu, dan nilai ini menjadi dasar dalam perhitungan *Gain*. Ringkasan hasil ini digunakan untuk melanjutkan proses perhitungan gain pada langkah berikutnya.

Berdasarkan hasil tersebut, nilai *Gain* untuk atribut Menu dihitung menggunakan rumus (3) sebagai langkah untuk menentukan pengurangan ketidakpastian yang dihasilkan oleh atribut ini. Perhitungan *Gain* diperlukan untuk menilai kontribusi atribut dalam memisahkan data berdasarkan kelas. Semakin besar nilai *Gain*, semakin besar kemampuan atribut tersebut dalam mengurangi ketidakpastian. Gambar 4 menyajikan hasil perhitungan *Gain* yang menjadi dasar dalam proses pemilihan atribut terbaik pada algoritma C4.5.

Rumus untuk menghitung nilai *gain*:

$$Gain\ (S, A) = Entropy\ (S) - \sum_{i=1}^{n} \frac{|S_j|}{|S|} \times Entropy\ (S_i) \quad (3)$$

$$Gain(Menu) = 1.3704 - 0 = 1.3704$$

Atribut	Entropy(S)	Entropy Split	Gain
Menu	1.3704	0	1.3704

Gambar 4. Hasil Tabel Perhitungan Gain

Setelah itu dilakukan perhitungan *split information* pada atribut Menu dengan menghitung proporsi masing-masing nilai atribut terhadap keseluruhan data menggunakan rumus (3). Tahap ini dilakukan karena *split information* berfungsi untuk menormalkan nilai *Gain* sehingga atribut dengan kategori sangat banyak tidak otomatis menjadi atribut terbaik. Nilai *split information* akan digunakan dalam perhitungan *Gain ratio*. Hasil perhitungan lengkap ditunjukkan pada Gambar 5.

Rumus untuk menghitung nilai *split information*:

$$Splitinfo\ (S, A) = -\sum_{j=1}^{k} \frac{s_j}{s}\ log^2 \frac{s_j}{s} \quad (4)$$

Proporsi:
1/5 = 0.2 (3 kategori)
2/5 = 0.4 (Lele Goreng)
$SI = -[3(0.2log_2 0.2) + 0.4log_2 0.4]$

Hitung log:
$0.2log_2 0.2 = -0.464$
$0.4log_2 0.4 = -0.529$
$SI = -[3(-0.464) + (-0.529)]$

turnitin Page 12 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

turnitin Page 13 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

p-ISSN : 2803-1507 e-ISSN : 2803-1531

$SI = -[-1.392 - 0.529]$
$SI = 1.921$

Nilai Menu	Proporsi	-p log2 p
Ayam Bakar	0.2	0.464
Lele Bakar	0.2	0.464
Ayam Goreng	0.2	0.464
Lele Goreng	0.4	0.529
Total		1.921

Gambar 5. Hasil Tabel Perhitungan Split Information

Sebagaimana terlihat pada Gambar 6, nilai *Gain Ratio* diperoleh dari perbandingan antara *Gain* dan *Split Information* menggunakan rumus (4). Langkah ini bertujuan untuk memilih atribut yang tidak hanya memiliki nilai *Gain* tinggi, tetapi juga memiliki tingkat pembagian kategori yang proporsional. Perhitungan *Gain Ratio* diperlukan untuk menghindari bias terhadap atribut dengan banyak kategori. Hasil perhitungan *Gain Ratio* ditunjukkan pada Gambar 6.

$$GainRatio\ (S, A) = \frac{\text{Gain(S, A)}}{\text{Split(S, i)}} \quad (5)$$

$$GainRatio = \frac{1.3704}{1.921} = 0.713$$

Atribut	Entropy Awal	Entropy Setelah Split	Gain	Split Info	Gain Ratio
Menu	1.3704	0	1.3704	1.921	0.713

Gambar 6. Hasil Tabel Perhitungan Gain Ratio

Perhitungan berikut hanya digunakan sebagai contoh penerapan algoritma *Decision Tree* C4.5 untuk menjelaskan proses perhitungan *entropy*, *Gain*, *split information*, dan *Gain Ratio*. Berdasarkan hasil perhitungan, atribut Menu memiliki nilai *Gain Ratio* tertinggi sehingga pada contoh ini atribut Menu terpilih sebagai node awal. Perhitungan ini menunjukkan bagaimana algoritma menentukan atribut terbaik sebelum membangun struktur pohon keputusan.

3.2.4.1. *Apply Model*

Apply model merupakan tahap penerapan model *Decision Tree* yang telah dibangun menggunakan data pelatihan. Pada tahap ini, model digunakan untuk memprediksi kelas atau kategori dari data baru berdasarkan pola yang telah dipelajari sebelumnya. Proses ini melibatkan input data yang sesuai dengan atribut yang digunakan saat membangun model, kemudian model menghasilkan keputusan berupa hasil klasifikasi [24]. Dalam penelitian ini, model digunakan untuk memprediksi menu yang berpotensi menjadi *best seller* di Warung Makan Lalapan Haris, seperti ayam goreng, ayam bakar, lele goreng, atau lele bakar. Hasil prediksi tersebut dapat digunakan pemilik warung untuk merencanakan persediaan bahan baku dan strategi penjualan [25].

3.2.4.2. *Performance Classification*

Performance Classification digunakan untuk menilai seberapa baik model klasifikasi dalam memprediksi kelas data. Evaluasi biasanya menggunakan *Confusion Matrix* dan metrik seperti akurasi, presisi, *recall* dan *F1-score*. Metrik-metrik ini membantu melihat ketepatan, kemampuan mengenali kelas, serta keseimbangan performa model, sehingga efektivitas *Decision Tree* dapat dinilai secara menyeluruh [26]. Metrik kunci yang didefinisikan di dalam matriks kebingungan dan untuk mengevaluasi performa model yang dilatih, nilai metrik performa dapat dihitung menggunakan persamaan berikut:

Model klasifikasi perlu dievaluasi menggunakan akurasi untuk mengetahui seberapa besar proporsi prediksi yang benar terhadap seluruh data pengujian. Nilai akurasi tersebut dihitung menggunakan rumus (6).

$$Accurancy = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \quad (6)$$

turnitin Page 13 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

turnitin Page 14 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

p-ISSN : 2803-1507 e-ISSN : 2803-1531

Nilai yang dihasilkan dari rumus ini menunjukkan tingkat ketepatan keseluruhan model dalam melakukan klasifikasi. Meskipun demikian, akurasi perlu diinterpretasikan secara hati-hati terutama pada dataset dengan distribusi kelas yang tidak seimbang.

Precision digunakan untuk mengukur seberapa tepat model dalam menghasilkan prediksi positif yang benar. Nilai *precision* dihitung menggunakan rumus (7).

$$Precision = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (7)$$

Nilai *precision* yang tinggi menunjukkan bahwa model jarang memberikan prediksi positif yang keliru. Metrik ini penting terutama pada kasus ketika kesalahan positif dapat menyebabkan dampak signifikan.

Recall digunakan untuk mengetahui kemampuan model dalam menemukan seluruh data yang termasuk dalam kelas positif. Nilai *recall* dihitung menggunakan rumus (8).

$$Recall = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (8)$$

Nilai *recall* yang tinggi menunjukkan bahwa model tidak banyak melewatkan data positif. Metrik ini sangat penting ketika tujuan analisis adalah meminimalkan jumlah data relevan yang tidak terdeteksi.

F1-score merupakan metrik yang menggabungkan *precision* dan *recall* untuk memberikan penilaian yang lebih seimbang. Nilai *F1-score* dihitung menggunakan rumus (9).

$$F1 - Score = \frac{2x(Precision\ x\ recall)}{Precision + Recall} \quad (9)$$

Metrik ini memberikan gambaran performa model yang lebih stabil ketika terjadi ketidakseimbangan data antar kelas. *F1-score* sangat bermanfaat ketika *precision* dan *recall* perlu dinilai secara bersamaan.

3.2.5 *Interpretation / Evaluation* (Interpretasi dan Evaluasi)

Tahap akhir dalam proses KDD adalah melakukan evaluasi terhadap model serta menafsirkan pengetahuan yang berhasil diperoleh. Pada penelitian ini, evaluasi model *Decision Tree* dilakukan di RapidMiner dengan menggunakan sejumlah metrik seperti akurasi, *precision, recall, dan confusion matrix* untuk menilai kualitas kinerja model [27]. Selain itu, struktur pohon keputusan dan aturan *(rules*) yang dihasilkan dianalisis lebih lanjut guna mengidentifikasi atribut yang paling berpengaruh dalam menentukan status menu. Hasil interpretasi ini menjadi dasar untuk memahami pola keputusan yang terbentuk dan memastikan bahwa model tidak hanya akurat, tetapi juga relevan secara praktis.

4. HASIL DAN PEMBAHASAN

4.1. *Hasil Pengolahan Dataset*

Dataset yang digunakan terdiri dari 500 data *record* penjualan menu yang berisi empat atribut, yaitu *Menu*, *Jumlah_Pelanggan*, *Jumlah_Terjual*, dan *Status_Menu*. Kolom Status_Menu merupakan kelas yang akan diprediksi menggunakan Algoritma *Decision Tree* C4.5. Sebagian data ditunjukkan pada Gambar 7. Berdasarkan data tersebut terlihat bahwa tiap menu memiliki variasi jumlah pelanggan dan jumlah terjual yang berbeda, sehingga cocok digunakan sebagai data klasifikasi untuk menentukan kategori: *Best Seller, Medium Seller, Low Seller*

turnitin Page 14 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

turnitin Page 15 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

p-ISSN : 2803-1507 e-ISSN : 2803-1531

Row No.	Menu	Jumlah_Pelanggan	Jumlah_Terjual	Status_Menu
1	Lele Goreng	15	15	Low Seller
2	Lele Goreng	13	18	Low Seller
3	Lele Goreng	18	27	Low Seller
4	Ayam Bakar	18	40	Medium Seller
5	Lele Bakar	20	20	Low Seller
6	Ayam Bakar	15	28	Medium Seller
7	Ayam Bakar	17	36	Medium Seller
8	Lele Goreng	9	14	Low Seller
9	Lele Goreng	7	8	Low Seller
10	Ayam Goreng	15	51	Best Seller
11	Lele Goreng	16	23	Low Seller
12	Lele Bakar	17	25	Low Seller
13	Lele Bakar	16	17	Low Seller
14	Lele Bakar	20	21	Low Seller
15	Ayam Bakar	5	11	Medium Seller
16	Lele Goreng	6	6	Low Seller

Gambar 7. Hasil Pengolahan Dataset

4.2. *Alur Proses (Workflow)*

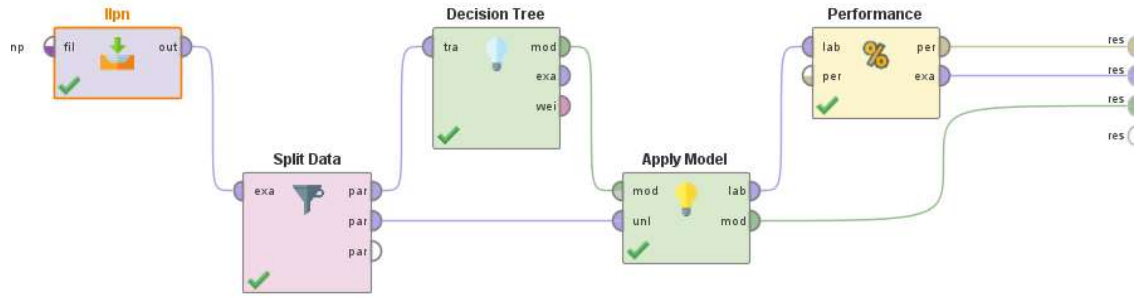
Proses pemodelan menggunakan RapidMiner dimulai dengan memasukkan dataset melalui operator *Read/Import* Data, yang berfungsi membaca seluruh atribut seperti Menu, Jumlah_Pelanggan, Jumlah_Terjual, dan Status_Menu. Data yang telah dimuat kemudian dibagi menggunakan operator *Split Data* dengan proporsi 70% sebagai data pelatihan dan 30% sebagai data pengujian. Pembagian ini dilakukan untuk memastikan bahwa model *Decision Tree* dilatih pada sebagian data dan dievaluasi menggunakan data baru yang belum pernah dilihat oleh model. Selanjutnya, bagian training dialirkan ke operator Decision Tree, di mana algoritma membangun struktur pohon keputusan berdasarkan perhitungan *entropy, information gain, split information, dan gain ratio* untuk menentukan atribut terbaik pada setiap pemecahan cabang. Setelah model selesai dibangun, operator *Apply Model* digunakan untuk menerapkan model tersebut pada data testing sehingga menghasilkan label prediksi untuk setiap *record*.

Tahap akhir dilakukan oleh operator P*erformance* yang menghitung nilai akurasi, *precision, recall*, serta menghasilkan *confusion matrix* sebagai dasar untuk mengevaluasi kemampuan model dalam mengklasifikasikan menu ke dalam kategori *Low Seller, Medium Seller,* atau *Best Seller*. Proses keseluruhan ini memberikan gambaran lengkap tentang bagaimana model *Decision Tree* dibangun, diuji, dan dievaluasi di RapidMiner. Selain itu, hasil evaluasi ini memberikan informasi penting mengenai efektivitas model dalam mendukung proses pengambilan keputusan bisnis. Gambar 8 menyajikan *workflow* RapidMiner yang digunakan dalam penelitian ini dalam membentuk model *Decision Tree*.

turnitin Page 15 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

turnitin Page 16 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

p-ISSN : 2803-1507 e-ISSN : 2803-1531



Gambar. 8 Alur proses

4.3. *Hasil Pembentukan Pohon Keputusan (Algoritma Decision Tree C4.5)*

Pohon keputusan pada Gambar 9 merupakan hasil proses Algoritma *Decision Tree* C4.5 dalam memprediksi Status_Menu berdasarkan atribut yang tersedia pada dataset. Struktur pohon yang dihasilkan memperlihatkan bagaimana setiap menu dipetakan ke dalam kategori tertentu sesuai pola penjualannya. Pada bagian yang mengarah ke kategori *Best Seller*, algoritma secara langsung menempatkan Ayam Goreng sebagai menu yang memiliki performa penjualan paling tinggi. Keputusan ini muncul karena dalam dataset, Ayam Goreng menunjukkan konsistensi penjualan yang jauh lebih unggul dibandingkan menu lainnya. Sementara itu, jalur yang mengarah pada kategori *Medium Seller* menunjukkan bahwa Ayam Bakar menjadi menu yang sesuai dengan karakteristik penjualan tingkat menengah. Pola penjualan Ayam Bakar yang stabil menyebabkan algoritma tidak memerlukan pemisahan tambahan dan langsung memasukkannya ke dalam kategori ini.

Bagian pohon yang menggambarkan kelas *Low Seller* memiliki struktur percabangan yang lebih kompleks karena sangat dipengaruhi oleh nilai Jumlah_Pelanggan. Ketika jumlah pelanggan melebihi 3.500, percabangan lanjutan digunakan untuk menentukan apakah menu tersebut termasuk Lele Goreng atau Lele Bakar. Jika jumlah pelanggan berada di atas 21.500, keputusan cenderung mengarah pada Lele Goreng. Dalam kondisi jumlah pelanggan lebih rendah dari batas tersebut namun tetap di atas 4.500, hasilnya tetap mengarah pada Lele Goreng. Namun ketika nilainya berada di bawah 4.500, klasifikasi berubah menjadi Lele Bakar. Pada situasi jumlah pelanggan kurang dari atau sama dengan 3.500, model secara konsisten mengelompokkan menu tersebut ke dalam kategori Lele Bakar.

Penjelasan ini menunjukkan bahwa Lele Goreng dan Lele Bakar merupakan menu yang paling dominan dalam kategori *Low Seller*. Jumlah pelanggan menjadi variabel paling berpengaruh dalam proses pemilahan untuk kategori ini, sehingga menghasilkan pembentukan aturan yang lebih rinci dibandingkan kategori lainnya. Variasi nilai pada atribut Jumlah_Pelanggan memberikan bentuk percabangan yang berbeda sehingga setiap kategori dapat dipetakan sesuai pola penjualannya. Dengan demikian, pohon keputusan pada Gambar 9 menggambarkan struktur klasifikasi yang logis berdasarkan pola penjualan aktual dalam dataset.



Gambar 9. Hasil Pembentukan Pohon Keputusan

turnitin Page 16 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

turnitin Page 17 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

p-ISSN : 2803-1507 e-ISSN : 2803-1531

4.4. *Hasil Evaluasi Model (Confusion Matrix)*

Tingkat akurasi model mencapai 84%, yang menunjukkan bahwa sebagian besar data uji dapat diklasifikasikan dengan benar oleh algoritma C4.5. Secara umum, model mampu mengenali pola penjualan setiap menu, meskipun masih terdapat perbedaan performa antar kelas. Kategori Lele Goreng memiliki *precision* 55,56% dan *recall* 100%; seluruh data kelas ini berhasil terdeteksi (*recall* tinggi), tetapi rendahnya *precision* mengindikasikan banyak prediksi yang keliru diarahkan ke kelas tersebut. Sebaliknya, Ayam Bakar serta Ayam Goreng menunjukkan *precision* dan *recall* sempurna, yang menggambarkan bahwa pola kedua menu ini sangat jelas dan mudah dipelajari oleh model.

Pada kelas Lele Bakar, *precision* mencapai 100% tetapi *recall* hanya 20%. Artinya, setiap prediksi Lele Bakar selalu benar, namun sebagian besar data sebenarnya tidak teridentifikasi dan justru diklasifikasikan sebagai menu lain. Kondisi ini biasanya terjadi ketika pola data kelas minor tidak cukup berbeda atau jumlah datanya terlalu sedikit, sehingga model tidak memperoleh cukup contoh untuk mempelajari batasan kelas secara tegas. Rendahnya performa pada Lele Bakar terutama disebabkan kemiripan nilai atribut dengan menu lain, sehingga pohon keputusan kesulitan membuat pemisahan yang jelas. Selain itu, atribut yang digunakan belum cukup informatif untuk membedakan menu yang pola penjualannya berdekatan. Dampaknya juga terlihat pada Lele Goreng, di mana *recall* tinggi tetapi *precision* rendah indikasi model sering memilih kelas tersebut saat ragu.

Secara keseluruhan, hasil evaluasi menunjukkan adanya ketidakseimbangan data (*imbalance*). Kelas dengan jumlah data besar seperti Ayam Goreng dan Ayam Bakar tampil sangat dominan, sementara kelas kecil seperti Lele Bakar tidak memperoleh representasi yang memadai selama pelatihan. Karena algoritma C4.5 cenderung memilih *split* yang memberikan *information gain* terbesar yang biasanya berasal dari kelas mayoritas kelas minor akhirnya kurang terwakili dalam aturan pohon keputusan. Distribusi nilai precision dan recall pada setiap kelas menggambarkan bagaimana model mengenali pola yang berbeda untuk kelas mayoritas dan minor secara terpisah. Dengan demikian, perbedaan performa antar kelas sangat dipengaruhi oleh tumpang tindih pola atribut, kurangnya fitur pembeda, dan distribusi data yang tidak seimbang, sebagaimana ditunjukkan pada Gambar 10.

accuracy: 84.00%

	true Lele Goreng	true Ayam Bakar	true Lele Bakar	true Ayam Goreng	class precision
pred. Lele Goreng	5	0	4	0	55.56%
pred. Ayam Bakar	0	11	0	0	100.00%
pred. Lele Bakar	0	0	1	0	100.00%
pred. Ayam Goreng	0	0	0	4	100.00%
class recall	100.00%	100.00%	20.00%	100.00%	

Gambar 10. Hasil Evaluasi Model

4.5. *Struktur Pohon Keputusan Lengkap (Text Tree)*

Bagian ini memperlihatkan bagaimana pohon dimulai dari atribut Status_Menu sebagai akar, kemudian bercabang mengikuti aturan-aturan yang dibentuk oleh Algoritma *Decision Tree* C4.5. Pada cabang tertentu, Ayam Goreng dan Ayam Bakar langsung ditempatkan pada kategorinya masing-masing karena pola penjualannya sudah cukup jelas sehingga tidak memerlukan pemisahan tambahan. Algoritma membangun struktur pohon berdasarkan atribut yang memiliki *information gain* tertinggi, sehingga setiap percabangan mencerminkan pola distribusi data asli. Setiap node pada pohon mengandung jumlah kasus aktual dari dataset sehingga memudahkan interpretasi kontribusi atribut terhadap klasifikasi.

Berbeda dengan itu, Lele Goreng dan Lele Bakar melalui proses pemilahan yang lebih panjang. Percabangan yang terbentuk menunjukkan adanya beberapa aturan bertahap, seperti keputusan berdasarkan nilai Jumlah_Pelanggan yang berada di atas 3.500, kemudian bercabang lagi pada batas 21.500, dilanjutkan dengan ambang lain seperti 4.500, dan seterusnya. Rangkaian aturan tersebut memperlihatkan bahwa atribut Jumlah_Pelanggan menjadi faktor yang paling dominan digunakan model dalam menentukan apakah sebuah menu termasuk ke dalam kategori Lele Goreng atau Lele Bakar. Struktur pohon keputusan lengkap ditunjukkan pada coding tree berikut, yang menampilkan setiap node, percabangan, dan jumlah kasus aktual pada dataset:

turnitin Page 17 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

turnitin Page 18 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

p-ISSN : 2803-1507 e-ISSN : 2803-1531

Tree

```
Status_Menu = Best Seller: Ayam Goreng {Lele Goreng=0, Ayam Bakar=0, Lele Bakar=0, Ayam Goreng=107}
Status_Menu = Low Seller
|   Jumlah_Pelanggan > 3.500
|   |   Jumlah_Pelanggan > 21.500: Lele Goreng {Lele Goreng=2, Ayam Bakar=0, Lele Bakar=0, Ayam Goreng=0}
|   |   Jumlah_Pelanggan ≤ 21.500
|   |   |   Jumlah_Pelanggan > 4.500: Lele Goreng {Lele Goreng=114, Ayam Bakar=0, Lele Bakar=108, Ayam Goreng=0}
|   |   |   Jumlah_Pelanggan ≤ 4.500: Lele Bakar {Lele Goreng=1, Ayam Bakar=0, Lele Bakar=3, Ayam Goreng=0}
|   Jumlah_Pelanggan ≤ 3.500: Lele Bakar {Lele Goreng=0, Ayam Bakar=0, Lele Bakar=6, Ayam Goreng=0}
Status_Menu = Medium Seller: Ayam Bakar {Lele Goreng=0, Ayam Bakar=134, Lele Bakar=0, Ayam Goreng=0}
```

4.6. *Tabel Pendukung Penelitian (Penjualan Menu)*

Penjualan menu dianalisis berdasarkan atribut persentase penjualan, status menu, dan peringkat penjualan. Setiap menu dicatat jumlah persentasenya sehingga dapat dibandingkan secara kuantitatif. Data ini menunjukkan variasi penjualan antara menu goreng dan menu lele. Ringkasan data tersebut disajikan dalam Tabel 2, yang mempermudah pembaca memahami performa tiap menu.

Tabel 2. Tabel Pendukung Penelitian

Menu	Persentase Penjualan(%)	Status Menu	Rangking Penjualan
Ayam Goreng	40%	Best Seller	1
Ayam Bakar	24%	Medium Seller	2
Lele Goreng	20%	Low Seller	3
Lele Bakar	16%	Low Seller	4

Tabel 2 menyajikan ringkasan penjualan setiap menu berdasarkan persentase penjualan, status menu, dan ranking. Dari tabel terlihat bahwa Ayam Goreng menempati posisi pertama dengan persentase penjualan 40%, sehingga dikategorikan sebagai *Best Seller*. Menu ini menjadi pilihan utama pelanggan, menunjukkan preferensi yang kuat terhadap menu goreng. Ayam Bakar berada di peringkat kedua (24%, Cukup Laku), menandakan menu ini masih diminati, namun tidak setinggi Ayam Goreng. Sementara itu, Lele Goreng (20%) dan Lele Bakar (16%) berada di peringkat ketiga dan keempat, termasuk kategori Kurang Laku, menunjukkan rendahnya minat pelanggan terhadap menu lele. Tabel ini membantu memperjelas preferensi pelanggan, yaitu menu yang paling diminati dan menu yang kurang diminati, sehingga menjadi dasar pengelolaan menu di warung untuk meningkatkan penjualan menu yang kurang populer dan memaksimalkan potensi menu yang sudah diminati.

5. DISKUSI

Hasil penelitian menunjukkan bahwa model klasifikasi C4.5 mampu mengidentifikasi pola penjualan secara konsisten, terutama untuk menu Ayam Goreng dan Ayam Bakar yang memiliki performa stabil sebagaimana juga dijelaskan dalam [5]. Temuan ini mengindikasikan bahwa atribut jumlah pelanggan dan jumlah penjualan berperan kuat dalam membentuk struktur *decision tree*. Kestabilan klasifikasi pada dua menu tersebut serupa dengan kecenderungan pola yang dijelaskan pada [7] terkait pengaruh atribut dominan dalam pembentukan keputusan model. Temuan ini memperlihatkan bahwa struktur C4.5 dapat mengikuti pola data yang memiliki konsistensi penjualan yang tinggi.

Perbandingan dengan penelitian sebelumnya memperlihatkan bahwa algoritma C4.5 cenderung lebih mudah mengenali kelas mayoritas, suatu kondisi yang juga diuraikan dalam [12] saat membahas kecenderungan model terhadap distribusi kelas. Performanya yang menurun pada menu Lele Goreng dan Lele Bakar menunjukkan adanya tumpang tindih atribut, situasi yang juga dicatat pada penelitian lain mengenai sensitivitas algoritma terhadap kelas minor [18]. Perbedaan stabilitas klasifikasi ini memberikan gambaran bahwa ketidakseimbangan kelas memengaruhi akurasi model pada kategori Low Seller. Temuan ini menunjukkan adanya pola dinamis pada data yang tidak selalu selaras dengan distribusi yang lebih sederhana dalam penelitian lain.

Hasil penelitian memperlihatkan bahwa kategori Low Seller memerlukan percabangan yang lebih kompleks karena atribut jumlah pelanggan menjadi faktor dominan dalam pembentukan cabang model. Pola tersebut memiliki kemiripan dengan bahasan dalam [14] yang menyebutkan bahwa atribut numerik dapat menghasilkan pemisahan berlapis ketika distribusi datanya tidak stabil. Performa model pada kondisi ini menunjukkan bahwa C4.5 masih dapat mengikuti pola meskipun struktur datanya lebih bervariasi. Temuan tersebut memperlihatkan bahwa karakteristik data memiliki peran besar dalam menentukan bentuk keputusan algoritma, terutama pada kategori dengan pola penjualan yang fluktuatif.

Implikasi praktis dari temuan ini menunjukkan bahwa model C4.5 dapat dimanfaatkan untuk membantu pengelola usaha kuliner dalam mengidentifikasi kategori performa menu secara lebih sistematis. Informasi

turnitin Page 18 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

turnitin Page 19 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

p-ISSN : 2803-1507 e-ISSN : 2803-1531

yang dihasilkan model dapat mendukung penataan strategi operasional seperti pengelolaan persediaan, prioritas produksi, dan penyusunan menu pada daftar hidangan. Secara teoritis, pola yang terbentuk dari hasil klasifikasi memberikan tambahan pemahaman mengenai cara algoritma *decision tree* memodelkan penjualan dalam konteks usaha kuliner skala kecil. Meskipun demikian, penelitian ini memiliki keterbatasan pada variasi atribut dan ketidakseimbangan data sehingga interpretasi hasil perlu disesuaikan dengan ruang lingkup penelitian.

6. KESIMPULAN DAN SARAN

6.1. *Kesimpulan*

Penelitian ini menghasilkan tiga kesimpulan utama terkait karakteristik penjualan setiap kategori menu berdasarkan model klasifikasi C4.5 yang dibangun dari 500 data penjualan Warung Makan Lalapan Haris. Pertama, kategori *Best Seller* secara konsisten ditempati oleh Ayam Goreng, yang memiliki volume penjualan tinggi dan stabil sehingga mudah dikenali model dengan akurasi yang sangat baik. Kedua, kategori *Medium Seller* diisi oleh Ayam Bakar, yang menunjukkan performa penjualan menengah namun relatif konsisten, sehingga pola klasifikasinya lebih sederhana tanpa memerlukan percabangan atribut yang kompleks. Ketiga, kategori *Low Seller*, yang terdiri dari Lele Goreng dan Lele Bakar, memiliki pola penjualan fluktuatif dan jumlah pelanggan yang rendah, sehingga model memerlukan beberapa lapisan aturan berbasis atribut Jumlah_Pelanggan untuk mencapai ketepatan klasifikasi.

Kinerja model menunjukkan akurasi 84% dengan performa sangat baik pada Ayam Goreng dan Ayam Bakar. Pada sisi lain, Lele Goreng memiliki *precision* 55,56% dan *recall* 100%, sedangkan Lele Bakar memperoleh *precision* 100% namun *recall* hanya 20%, menjadikannya kelas yang paling sulit dikenali. Perbedaan performa ini menggambarkan tingkat kompleksitas yang berbeda antar menu dalam proses pembentukan pohon keputusan. Secara praktis, temuan ini menunjukkan bahwa model dapat digunakan sebagai alat bantu pengambilan keputusan, khususnya dalam perencanaan stok bahan baku, penentuan strategi promosi, serta pemahaman pola permintaan harian. Kategori *Low Seller* secara khusus dapat menjadi fokus peningkatan strategi pemasaran untuk mengurangi risiko pemborosan persediaan.

6.2. *Saran*

Berdasarkan hasil penelitian, disarankan agar pemilik Warung Makan Lalapan Haris memanfaatkan informasi dari model klasifikasi ini dalam merencanakan kebutuhan bahan baku harian serta mengoptimalkan pengelolaan menu. Khusus untuk menu yang tergolong *low seller*, pemilik dapat menerapkan strategi promosi seperti memberikan potongan harga, membuat paket bundling, atau menampilkan menu tersebut pada posisi yang lebih strategis dalam daftar menu untuk meningkatkan daya tarik bagi pelanggan. Selain itu, pemilik juga dapat melakukan evaluasi berkala terhadap cita rasa, penyajian, maupun harga untuk mengetahui faktor penyebab rendahnya penjualan, sehingga langkah perbaikan dapat dilakukan secara lebih tepat sasaran. Penelitian di masa mendatang disarankan untuk memasukkan atribut tambahan yang berpotensi memengaruhi pola penjualan, seperti waktu transaksi, hari operasional, serta faktor promosi, sehingga model dapat mempelajari pola yang lebih detail dan meningkatkan akurasi prediksi. Selain itu, evaluasi komparatif terhadap berbagai algoritma klasifikasi misalnya *Random Forest, Naive Bayes,* atau *K-Nearest Neighbor*, perlu dilakukan agar diperoleh model yang paling tepat dan stabil dalam memprediksi menu terlaris.

DAFTAR PUSTAKA

[1] M. Ferdyandi, N. Y. Setiawan, and F. A. Bachtiar, "Prediksi Potensi Penjualan Makanan Beku berdasarkan Ulasan Pengguna Shopee menggunakan Metode Decision Tree Algoritma C4.5 dan Random Forest (Studi Kasus Dapur Lilis)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 6, no. 2, pp. 588–596, 2022, [Online]. Available: https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/10560

[2] H. Dhika *et al.*, "Penerapan Algoritma Decision Tree Dalam Penjualan Handphone," *J. Rekayasa Komputasi Terap.*, vol. 1, no. 4, pp. 239–246, 2021, doi: 10.30998/jrkt.v1i04.6157.

[3] F. Rahmadayanti, A. Asminah, and D. T. Cahaya, "Penerapan Algoritma C4.5 pada Produk Penjualan Makanan Ringan," *J. Teknol. Inf. MURA*, vol. 16, no. 1, pp. 62–69, 2024, doi: 10.32767/jti.v16i1.2294.

[4] M. H. S. Wibowo, M. D. Al Ayubi, and E. Rilvani, "Klasifikasi Volatilitas Harga Daging Ayam dan Cabe Rawit Merah dengan Decision Tree," *J. Komput. dan Teknol.*, vol. 4, no. 2, pp. 138–153, 2025,

turnitin Page 19 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

turnitin Page 20 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

p-ISSN : 2803-1507 e-ISSN : 2803-1531

doi: 10.64626/jukomtek.v4i2.455.

[5] A. Ardhy, T. Maulana, and D. Arman, "Penerapan Metode Decision Tree C4.5 untuk Klasifikasi Data Kandidat Tenaga Kerja pada Perusahaan Outsourcing," *J. Apl. Sains, Informasi, Elektron. dan Komput.*, vol. 6, no. 1, pp. 41–48, 2024, doi: 10.26905/jasiek.v6i1.12670.

[6] Wirawan, D. Alan, Rohman, and Abdul, "Consumer Satisfaction Analysis at Inos Coffee & Kitchen Using the C4.5 Algorithm: Analisis Kepuasan Konsumen Pada Inos Coffee & Kitchen Menggunakan Algoritma C4.5," *J. Inovtek Polbeng*, vol. 10, no. 1, pp. 526–537, 2025, doi: 10.35314/nbpz0h21.

[7] N. I. Yaman, A. R. Juwita, S. Arum, P. Lestari, and S. Faisal, "Perbandingan Kinerja Algoritma Decision Tree dan Random Forest untuk Klasifikasi Nutrisi pada Makanan Cepat Saji," *J. Algoritm.*, vol. 21, no. 2, pp. 184–195, 2024, doi: 10.33364/algoritma/v.21-2.1649.

[8] C. Supriadi, C. Sitinjak, D. Fujiama, N. D. Silalahi, S. Pertiwi, and A. Umar, "Enhancing Performance Using New Hybrid Intrusion Detection System," *J. Technol. Informatics Eng.*, vol. 1, no. 2, pp. 1–4, Jul. 2022, doi: 10.51903/jtie.v1i2.134.

[9] M. H. Haiedar and S. Kholifah, "The Impact of Occupational Safety, Workload, and Compensation on Employee Job Satisfaction in the Mining Industry: A Case Study of PT BCKA," *J. Manag. Informatics*, vol. 4, no. 1, pp. 587–598, Jan. 2025, doi: 10.51903/jmi.v4i1.53.

[10] A. Ramdani, C. D. Sofyan, F. Ramdani, M. Fauzi, A. Tama, and M. A. Rachmatsyah, "Algoritma Klasifikasi Data Mining untuk Memprediksi Masyarakat dalam Menerima Bantuan Sosial," *J. Ilm. Sist. Inf.*, vol. 1, no. 2, pp. 39–47, Jul. 2022, doi: 10.51903/juisi.v1i2.363.

[11] R. Rasendriya, A. O. Marundrury, Y. L. Jumadi, and A. D. Kuswanto, "Komparasi Decision Tree, Random Forest, dan k-NN Memprediksi Kelulusan Siswa Menggunakan Orange," *J. Komput. dan Teknol.*, vol. 4, no. 2, pp. 60–71, 2025, doi: 10.64626/jukomtek.v4i2.414.

[12] C. Aprianti, M. Faishal, and Y. Umaidah, "Penerapan Data Mining untuk Klasifikasi Penjualan Baju Muslim di Masa Pandemi Covid-19 Menggunakan Metode Algoritma C4.5 (Studi Kasus: Garaya Collection)," *J. Ilm. Wahana Pendidik.*, vol. 8, no. 1, pp. 104–112, 2022, doi: 10.5281/zenodo.5816231.

[13] R. yuninda Hutagalung and M. S. Rika Harman, S.Kom., "Analisis Kriteria Kamar Kos Layak Huni Menggunakan Algoritma C4.5 di Kota Batam," *J. Comasie (Computer Sci. Ind. Eng.*, vol. 8, no. 3, pp. 65–73, 2023, [Online]. Available: https://ejournal.upbatam.ac.id/index.php/comasiejournal/article/view/6935

[14] D. A. Pratama, I. R. Mutaqin, and K. R. Manuela, "Analisis Terjadinya Kanker Paru-Paru Pada Pasien Menggunakan Decision Tree: Penerapan Algoritma C4.5 Dan RapidMiner Untuk Menentukan Risiko Kanker Pada Gejala Pasien," *J. Tek. Mesin, Ind. Elektro dan Inform.*, vol. 2, no. 4, pp. 156–170, 2023, doi: 10.55606/jtmei.v2i4.3004.

[15] R. A. Saputra and A. Pratama, "Implementasi Decision Tree untuk Prediksi Harga Rumah di Daerah Tebet," *J. Inf. Syst. Manag.*, vol. 6, no. 2, pp. 4–10, 2025, doi: 10.24076/joism.2025v6i2.1928.

[16] N. N. Pradana, A. A. Subekti, and E. Rilvani, "Deteksi Transaksi Mencurigakan Menggunakan Decision Tree dan Logistic Regression dengan Mitigasi Ketidakseimbangan Kelas," *J. Media Akad.*, vol. 3, no. 8, pp. 3031–5220, 2025, doi: 10.62281/v3i8.2680.

[17] N. B. A. B. Arifin and A. Asmianto, "Sistem Prediksi Penjualan Menggunakan Kombinasi Metode Monte Carlo dan Decision Tree Berbasis Website," *MATHunesa J. Ilm. Mat.*, vol. 11, no. 2, pp. 274–286, 2023, doi: 10.26740/mathunesa.v11n2.p274-286.

[18] Z. Fatah and M. F. Wahid, "Penerapan Data Mining Menggunakan Algoritma C4.5 untuk Klasifikasi Penyakit Paru-Paru," *J. Ilm. Penelit. Mhs.*, vol. 3, no. 5, pp. 646–653, 2025, doi: 10.61722/jipm.v3i5.1513.

[19] R. P. Eshardiansyah, N. Sulistiyowati, and M. Jajuli, "Algoritma C4.5 Untuk Klasifikasi Jenis Kekerasan pada Anak (Kasus DP3A Kabupaten Karawang)," *J. Sains Komput. Inform.*, vol. 5, no. 2, pp. 687–696, 2021, doi: 10.30645/j-sakti.v5i2.368.

turnitin Page 20 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

turnitin Page 21 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177

p-ISSN : 2803-1507 e-ISSN : 2803-1531

[20] M. Rafi, N. Akbar, and E. K. Suni, "Penerapan Data Mining Untuk Memprediksi Penjualan Makanan Pada Kopi Krintji Menggunakan Algoritma Regresi Linear dan C4.5," *J. Ilmu Tek. dan Komput.*, vol. 6, no. 2, pp. 122–128, 2022, doi: 10.22441/jitkom.v6i2.005.

[21] A. Arif and A. Yulianto, "Klasifikasi Penjualan Makanan Hewan Peliharaan di Muezza Petshop01 Bogor Menggunakan Metode Algoritma C4.5," *REMIK Ris. dan E-Jurnal Manaj. Inform. Komput.*, vol. 6, no. 4, pp. 840–854, 2022, doi: 10.33395/remik.v6i4.11859.

[22] N. A. Lestari, "Penerapan Data Mining menggunakan Metode Decision Tree C4 . 5 untuk Prediksi Tingkat Kelulusan Mahasiswa (Studi Kasus : STMIK WIT)," Sekolah Tinggi Manajemen Informatika dan Komputer LIKMI Bandung, 2017. [Online]. Available: https://library.likmi.ac.id/show/80/pdf

[23] F. Arzalega and N. N. Pusparini, "Implementasi Data Mining pada Penjualan Main Course Pizza Menggunakan Algoritma C4.5 (Studi Kasus Pizza Hut Delivery Muara Karang Raya)," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 2, pp. 2283–2290, 2024, doi: 10.36040/jati.v8i2.9469.

[24] A. Ristyawan, A. Nugroho, and T. K. Amarya, "Optimasi Preprocessing Model Random Forest Untuk Prediksi Stroke," *J. Tek. Inform. dan Sist. Inf.*, vol. 12, no. 1, pp. 29–44, 2025, doi: 10.35957/jatisi.v12i1.9587.

[25] A. B. Prasetyo and I. H. Ikasari, "Penerapan Data Mining untuk Memprediksi Penjualan Kue Brownis Menggunakan Algoritma C4.5 pada Toko Amanda Brownis," *Log. J. Ilmu Komput. dan Pendidik.*, vol. 2, no. 1, pp. 292–298, 2023, doi: 10.1017/cbo9781107415324.004elian.

[26] A. Haryadi, A. N. P. Haryadi, W. Setiawan, and D. A. Fatah, "Penerapan Data Mining untuk Klasifikasi Penyakit Diabetes Menggunakan Metode Decision Tree," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 7, no. 6, pp. 1484–1495, 2024, doi: 10.32672/jnkti.v7i6.8145.

[27] S. O. Usnaat, Y. P. K. Kelen, and K. J. T. Seran, "Implementasi Metode Decision Tree dalam Memprediksi Penjualan Produk Terlaris (Studi Kasus: Kreatif Mart Atambua)," *Zo. J. Sist. Inf.*, vol. 7, no. 1, pp. 317–330, 2025, doi: 10.31849/zn.v7i1.23868.

turnitin Page 21 of 21 - Integrity Submission Submission ID trn:oid:::1:3442651177