

IMPLEMENTASI DATA MINING PADA PENJUALAN MAIN COURSE PIZZA MENGUNAKAN ALGORITMA C 4.5 (STUDI KASUS PIZZA HUT DELIVERY MUARA KARANG RAYA)

Fahmi Arzalega, Nur Nawaningtyas Pusparini

Teknik Informatika, Sekolah Tinggi Manajemen Informatika Dan Komputer Widuri
Jl. Palmerah Barat No. 353, RT.3/RW.5, Kel. Grogol Utara, Kec. Kebayoran Lama, Jakarta Selatan 12210
fahmia1941@kampuswiduri.ac.id

ABSTRAK

Ketika seorang kepala outlet *Pizza Hut Delivery* Muara Karang Raya sulit untuk menentukan dari *main course* pizza yang paling laris dan tidak laris. Selain itu kepala outlet *Pizza Hut Delivery* Muara Karang Raya belum mengetahui faktor apa yang bisa menentukan apakah produk *main course* *Pizza Hut Delivery* tersebut laris dan tidak laris. Tujuan dari penelitian ini adalah untuk mengetahui produk laris dan tidak laris pada *main course* *Pizza Hut Delivery* agar tepat sasaran dan faktor apa saja yang mempengaruhi algoritma C4.5 dalam menentukan produk laris dan tidak laris pada *main course*. Metode penelitian menggunakan Algoritma C4.5 dengan menghasilkan pohon keputusan. Pohon keputusan merupakan yang mempresentasikan aturan-aturan yang mudah dipahami untuk menemukan hubungan yang tersembunyi dari variabel input dan target. Pada penelitian ini akan menggunakan salah satu software data mining yaitu *Rapid Miner*. Dengan menggunakan aplikasi *Rapid Miner* hasil yang didapat pada sebuah hasil akurasi pada penelitian ini sebesar 88,24% dan penjualan *main course* laris pada *Pizza Hut Delivery* Muara Karang Raya adalah menu *American Favorite*, *Meat Lovers*, dan *Pepperoni*.

Kata kunci : Data Mining, Algoritma C4.5, Main Course, Rapid Miner

1. PENDAHULUAN

Makanan merupakan keadaan yang utama yang menjadi kebutuhan tiap manusia selama menjalankan kehidupan, maka dari itu akan munculnya beraneka ragam bisnis yang memasarkan macam jenis makanan. Namun, pada zaman sekarang makanan membuat sebuah perubahan dari “hanya sekedar untuk keperluan hidup” berganti sebagai sebuah “kultur atau gaya hidup”. Karena itu dapat dilihat sudah melimpah *owner* usaha yang bergerak dibidang bisnis makanan yang menawarkan menu yang sebelumnya normal-normal kemudian *owner* usaha bisnis makanan merubahnya menjadi sebuah makanan yang “modern dan kekinian” dan pada akhirnya membuat kesukaan bagi seluruh kalangan umur. Bahkan dalam masyarakat *modern*, orang selalu mencari makanan cepat saji yang praktis, yang menyebabkan munculnya beberapa bisnis yang menyajikan masakan jenis ini. Sektor industri makanan mendapat banyak keuntungan dari perkembangan informasi dan teknologi terbaru, yang mungkin dapat dikenal sebagai “kecerdasan bisnis”. Sebagai contohnya seperti penjualan makanan menggunakan teknik data mining [1].

Memanfaatkan *database* yang telah dimiliki semaksimal mungkin untuk mendapatkan fakta-fakta yang dapat digunakan untuk menentukan pilihan, keadaan ini bisa membuat munculnya suatu cabang pada keahlian yang baru untuk mengatasi masalah dalam mencari informasi-informasi atau sebuah pola yang memang penting atau menarik dari data yang jumlah sangat besar, yang disebut dengan data mining. Pada saat menggunakan teknik data mining berharap bisa memberikan banyak pengetahuan-pengetahuan baru yang sebelumnya tidak diketahui.

Algoritma C4.5 merupakan komponen dari algoritma pengelompokan Pohon Keputusan (*Decision Tree*). Data *training* digunakan sebagai input algoritma C4.5. Data *training* menggunakan data *field* dan sampel data yang akan digunakan untuk membuat pohon keputusan dan memilih label yang diinginkan. *Field – Field* data ini disebut sebagai atribut dan akan digunakan sebagai kriteria untuk melakukan klasifikasi data. Memilih atribut sebagai *root*, membuat cabang untuk setiap nilainya, memisahkan kasus menjadi cabang, dan mengulang prosedur untuk setiap cabang itu merupakan langkah – langkah dalam membangun algoritma ini. Analisis prediksi, data mining, dan text mining semuanya dapat dilakukan dengan algoritma C4.5 [2].

Pizza Hut Delivery didirikan pada tahun 2017 dibawah naungan PT.Sari Melati KencanaTbk, bisnis yang berspesialisasi dalam layanan pengiriman pizza ini dimulai pada tahun 1987 melalui waralaba yang dimiliki. *Pizza Hut Delivery* (PHD) merupakan restoran cepat saji (*Fast food*) pertama ada di Indonesia yang berfokus pada layanan pesan antar. *Pizza Hut Delivery* yang berkonsentrasi pada *delivery* ditujukan untuk menangkap segmen yang belum sepenuhnya tergarap oleh restoran *dine in*. *Pizza Hut Delivery* merupakan bisnis makanan yang mulai tumbuh agresif dari 23 outlet di tahun 2010, dalam 2 tahun *Pizza Hut Delivery* sukses mencapai 76.outlet ditahun 2012. Sejak itu menjadi pendorong pertumbuhan bisnis makanan dan per 31 Desember 2017 pangsa pasar *Pizza Hut Delivery* yang dilaporkan oleh Euromonitor tercatat 64,6% dengan sebanyak 156 outlet. Mengingat umat islam merupakan mayoritas penduduk Indonesia, menurut PT. Sari Melati Kencana Tbk, sangat penting

untuk memiliki sertifikat Halal. PT. Sari Melati Kencana Tbk mendapatkan sertifikasi halal dari MUI pada tahun 1997. Demi memberikan kenyamanan, keamanan, dan kebahagiaan pelanggan, PT. Sari Melati Kencana Tbk sampai saat ini telah mempertahankan sertifikat halal dengan cara tetap berkomitmen untuk selalu memproduksi produk yang halal.

Penulis melakukan riset di salah satu outlet *Pizza Hut Delivery* yang berlokasi di Muara Karang Raya bahwa ketersediaan data penjualan pada *main course* di *Pizza Hut Delivery* Muara Karang Raya pada data penjualannya belum memiliki cara menentukan produk laris dan tidak laris dengan tepat sasaran pada *main course* di PHD Muara Karang Raya.

Menurut wawancara yang dilakukan, penjualan di *Pizza Hut Delivery* Muara Karang Raya setiap harinya bisa terjual kurang lebih 200 loyang pan pizza, itu terbagi ke dalam banyak rasa dan juga ukuran yang ada di *main course* dan dalam sebulan kurang atau lebih 1500 loyang pan pizza. Untuk *main course* terdapat 12 item dan Ukuran Pizza terdapat 2 Ukuran yaitu Reguler dan Jumbo. Harga untuk *main course* dimulai dari Rp.90.000. Tapi dari data penjualan tersebut hanya di catat atau diinput saja setiap harinya tidak dimanfaatkan untuk mencari tahu kira-kira menu apa yang paling disukai konsumen, dengan adanya informasi ini bisa membantu dalam membuat strategi penjualan untuk kedepannya agar bisa meningkatkan sales penjualan. Mungkin dengan cara penjual melakukan sebuah inovasi baru terhadap menu-menu yang paling disukai konsumen.

Dari wawancara yang penulis lakukan terdapat suatu masalah yaitu ketika seorang kepala outlet *Pizza Hut Delivery* muara karang raya sulit untuk menentukan dari *main course* pizza yang paling laris dan tidak laris. Selain itu kepala outlet *Pizza Hut Delivery* muara karang raya belum mengetahui faktor apa saja yang bisa menentukan apakah produk *main course* *Pizza Hut Delivery* tersebut laris dan tidak laris.

Maka dari itu untuk permasalahan ini, penulis menerapkan teknik data mining *decision tree* algoritma C4.5 pada *Pizza Hut Delivery* muara karang. Dari hasil analisa menggunakan algoritma C4.5 diharapkan suatu informasi dari klasifikasi *main course* pizza yang laris dan tidak laris berdasarkan data set yang ada. Berdasarkan hasil analisa menggunakan algoritma C4.5 kepala outlet *Pizza Hut Delivery* muara karang raya dapat melakukan sebuah prediksi terhadap *main course* pizza dengan cara mengikuti trend dan kegemaran pelanggannya.

2. TINJAUAN PUSTAKA

Berikut ini adalah tinjauan pustaka yang akan dijelaskan yang terkait mengenai materi dan metode – metode yang akan dilakukan oleh penulis dalam melakukan penelitian ini.

2.1. Data Mining

Data mining merupakan prosedur menemukan pola dalam data yang memanfaatkan metode statistik, matematika, dan *machine learning*. Output akan menghasilkan sebuah informasi sangat penting dan bermanfaat. Lalu apa itu sebuah data?, sebuah data adalah sekumpulan fakta yang memberikan gambaran umum setiap kali mengumpulkan data maka data tersebut tersimpan dan pola dievaluasi secara manual dan teliti untuk menentukan apa yang akan terjadi [3].

Penerapan saintifik dipakai pada *Knowledge Discovery in Databases (KDD)*, yaitu data mining. Data mining adalah langkah dalam proses KDD dalam konteks ini. Proses menambang informasi bisnis yang berharga dari basis data yang sangat besar menggunakan teknik data mining yang dapat dianalogikan dengan mengekstraksi logam mulia dari sumbernya. Maka dari itu teknologi ini dapat digunakan untuk:

- Menciptakan ciri khas yang penting pada nilai atribut yang sedang digunakan.
- Menciptakan bentuk dan peranan yang berguna buat mendeskripsikan kelas atau rancangan data.
- Menciptakan satu jaringan terhadap nilai atribut dari seluruh data.

2.2. Algoritma C4.5

Algoritma C4.5 ialah teknik *decision Tree* yang memiliki algoritma klasifikasi data yang memiliki banyak kelebihan. Kelebihan ini termasuk kemampuan untuk memproses data numerik (*Continue*) dan diskret, kemampuan untuk menangani nilai atribut yang hilang, menghasilkan aturan yang mudah dipahami, dan menjadikan algoritma tercepat di antara yang lainnya [4].

Keakuratan prediksi model adalah kapasitasnya untuk memprediksi label kelas untuk data baru atau yang sebelumnya tidak diketahui. dalam hal seberapa cepat atau efisien model diproses selama pembuatan dan penggunaannya. Kapasitas model untuk secara akurat memprediksi data dengan nilai atribut yang hilang. Selain itu, skalabilitas, atau kapasitas untuk menghasilkan model secara efisien untuk sejumlah besar data, akan ditekankan. Bagian terakhir dari interpretabilitas adalah kemudahan pemahaman model.

Dalam membangun *decision Tree* dengan algoritma C4.5 hal yang pertama dilakukan adalah mencari akar dari atribut yang dimiliki. Lalu selanjutnya akan di buat sebuah cabang. Pada tahap selanjutnya mengulangi proses pada masing – masing cabang hingga seluruh masalah yang dimiliki cabang mempunyai kelas yang sama.

Entropy ialah sebuah perbandingan atau keberagaman. Pada saat melakukan penelitian data mining, *Entropy* merupakan bagian suatu alat mengukur atau parameter *heterogenitas* (keberagaman) pada suatu himpunan data yang dimiliki. Jika heterogen pada himpunan suatu data, maka semakin tinggi pula nilai yang dimiliki

entrophynya [5]. Pada dasar matematika memiliki rumus sebagai berikut :

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

Rumus diatas dapat dijelaskan sebagai berikut:

S = Nilai himpunan pada setiap kasus

N = Totalpartisi pada S

Pi = ProporsiSi terhadap S

Yang bisa dilihat pada nilai N ialah total nilai yang ada pada *attribut* target atau total kelas. Kemudian Pii merupakan pernyataan porsi atau rasio antar total sampel dikelas i dengan memiliki jumlah sampel pada himpunandata.

Gain ialah hasil informasi. Pada data *mining* *gainn* didefinisikan sebagai ukuran keefektian suatu *attribut* dalam mengklasifikasikan data. Secara matematika, *gainn* dirumuskan sebagai berikut:

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \quad (2)$$

Rumus diatas dapat dijelaskan sebagai berikut:

S = Kelompok pada data

A = Atribut pada data

N = Total nilai pada *attribut* A

|Si| = Total kasus pada partisi ke-i

|S| = Total kasus dalam S

Jika *attribut test* untuk node tersebut adalah *attribut* dengan nilai *gain* tertinggi. *Confusion matrix* ialah teknik yang biasa disebagai perhitungan presisi dalam konsep data mining. Informasi dalam matriks *confusion* diperlukan untuk menentukan efektivitas model klasifikasi. Menggabungkan informasi ini menjadi satu nilai digunakan untuk membandingkan kinerja model yang berbeda. Ini dapat dilakukan dengan menggunakan metrik kinerja [6].

Berikut ini adalah rumus untuk menghitung *confusion matrix*:

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (3)$$

Penjelasan dari rumus di atas adalah:

TP = TruePositive

TN = TrueNegative

FN = FalseNegative

FP= FalsePositif

2.3. Rapid Miner

Perusahaan *RapidMiner* menciptakan platform perangkat lunak sains data yang menawarkan rekomendasi untuk pembelajaran, ekstraksi data, dan analisis prediktif menggunakan pembelajaran mesin. Sebuah aplikasi yang mendukung tahapan yang berjumlah 24 ini masih berupa proses pembelajaran

yang meliputi rancangan data, hasil dan visualisasi, validasi, dan optimalisasi. *Rapidminer* ini biasanya dipakai untuk bisniskomersial serta penelitian, pendidikan, dan pengembangan aplikasi[7].

Perangkat lunak yang dapat diakses publik *RapidMiner* adalah solusi untuk analisi data *mining*, teks *mining*, dan analisis prediksi. Ini memberi pengguna wawasan melalui berbagai metode deskriptif dan prediksi sehingga mereka dapat membuat keputusan terbaik [8].

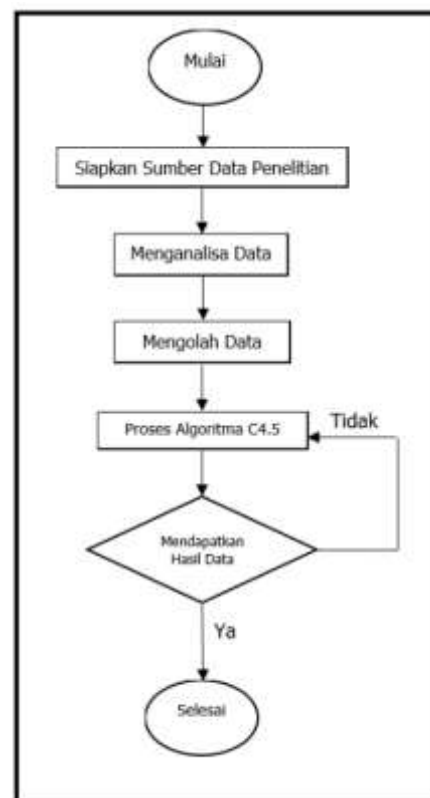
3. METODE PENELITIAN

Dalam melakukan penelitian ini penulis akan menjelaskan metodologi apa saja yang akan diterapkan pada penelitian ini. Dapat dijelaskan sebagai berikut.

3.1. Desain penelitian

Dalam menyusun dan melengkapi skripsi ini, penulis desain metode penelitian yang digunakan melalui tahapan – tahapan proses yang dilaksanakan dalam penentuan jenis data penelitian menggunakan metode algoritma C4.5 sebagai berikut:

1. Siapkan sumber data penelitian.
2. Menganalisa data.
3. Mengolah data.
4. Proses Algoritma C4.5
5. Mendapatkan hasil data.



Gambar 1. Desain penelitian

Pada gambar 1 dapat dijelaskan sebagai berikut, bahwa untuk memulai sebuah penelitian harus menyiapkan sebuah sumber data penelitian yang berguna untuk melakukan tahapan menganalisa data

yang berfungsi untuk menganalisa data yang bertujuan menghasilkan sebuah informasi atau menghasilkan pengetahuan dari data mentah. Setelah melakukan menganalisa data untuk melanjutkan sebuah penelitian harus melakukan proses pengolahan data yang dimana pengolahan adalah untuk melakukan sebuah proses pembersihan data, pengelompokan data berdasarkan karakteristik, dan melakukan *transformasi* data yang menjadikan sebuah data set. Selanjutnya setelah melakukan pengolahan data akan dituangkan hasil dari pengolahan data (*data set*) tersebut ke metode algoritma C4.5 dengan menggunakan aplikasi *rapid miner* dengan tujuan untuk menghasilkan sebuah pohon keputusan dan sebuah hasil produk laris dan tidak laris pada penelitian ini.

3.2. Teknik Pengumpulan Data

Selanjutnya akan dijelaskan lebih rinci mengenai teknik mendapatkan sebuah data yang dipakai dalam penelitian ini:

1. Wawancara

Wawancara yang dilaksanakan oleh penulis kepada pihak manajemen Pizza Hut Delivery Muara Karang Raya, yang berguna untuk mendapatkan informasi mengenai data penjualan pizza terkait untuk mengetahui menu main course pizza laris dan tidak laris.

2. Observasi

Observasi yang dilakukan oleh penulis dengan mengunjungi langsung tempat penelitian dalam hal ini di Pizza Hut Delivery Muara Karang Raya, untuk menentukan menu main course pizza yang laris dan tidak laris pada data penjualan.

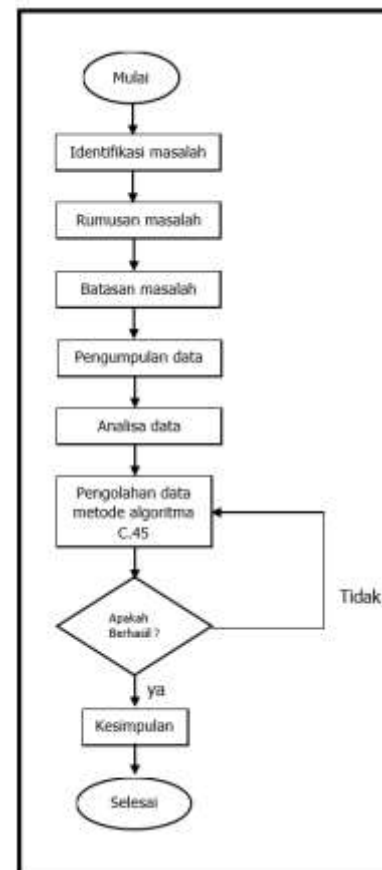
3. Studi Pustaka

Studi pustaka yang dilakukan oleh penulis dengan cara mencari referensi melalui jurnal terdahulu, buku, dan sebagainya yang berguna untuk mengumpulkan sumber data yang diinginkan pada saat melakukan penelitian ini.

Peneliti menyatukan semua seluruh teknik pengumpulan data yang telah disebutkan diatas dan diolah kembali untuk menghasilkan sebuah informasi dan data yang bermanfaat. Pada sumber data dan informasi yang bermanfaat ini akan dijadikan topik pembahasan pada penelitian ini.

3.3. Langkah – langkah penelitian

Pada tahapan ini penulis akan menjelaskan langkah-langkah proses tentang penerapan data mining algoritma C4.5 untuk melakukan prediksi laris dan tidak laris pada menu main course pada pizza dengan menggunakan aplikasi *Rapid.Miner*. Gambar dibawah adalah gambaran langkah – langkah yang dilaksanakan selama penelitian :



Gambar 2. Langkah penelitian

Berdasarkan gambar diatas bisa dijelaskan dibawah ini:

a. Identifikasi masalah

Untuk tahap awal yang dilakukan oleh penulis adalah mengenai identifikasi masalah yang ada dipenelitian berdasarkan apa yang terjadi disituasi dan kondisi sekarang, tetapi perlunya diketahui bahwa dalam menentukan suatu identifikasi masalah diperlukan sebuah sumber data yang valid dan konkrit agar tidak adanya timbul persepsi, oleh karna itu identifikasi masalah tersebut didapatkan dari hasil wawancara kepada manajemen outlet di *Pizza Hut Delivery* Muara Karang Raya.

b. Rumusan masalah

Pada tahapan ini penulis melakukan perumusan sebuah masalah dengan membuat sebuah pertanyaan mengenai masalah yang terjadi pada penelitian ini. Rumusan pada masalah itu akan menjadikan pengarah tujuan dari sebuah penelitian yang dilakukan.

c. Batasan masalah

Pada tahapan selanjutnya dalam melakukan sebuah batasan masalah pada penelitian akan berfokus pada permasalahan pada produk laris dan tidak laris pada menu *main course* pizza di *Pizza Hut Delivery* Muara Karang Raya. Dari permasalahan yang sudah didapatkan peneliti akan menentukan produk laris dan tidak laris pada menu *main course* pizza dan dari itu peneliti

akan menggunakan metode algoritma C4.5 pada data yang disediakan.

d. Pengumpulan data

Pada tahapan ini peneliti mengumpulkan data yang didapatkan dari data penjualan pertahun yang tersedia di outlet Pizza Hut Delivery Muara Karang Raya yang berupa file excel.

e. Analisa data

Pada tahapan ini peneliti akan menganalisa data terhadap apa saja yang dilakukan setelah mendapatkan data atau sebuah informasi berdasarkan penelitian pada data penjualan laris dan tidak laris pada menu *main course* pizza di *Pizza Hut Delivery* Muara Karang Raya menggunakan metode Algoritma C4.5, dari data tersebut dapat dianalisa dengan melakukan sebuah wawancara yang dilakukan pada manajemen outlet di *Pizza Hut Delivery* Muara Karang Raya. Lalu melakukan proses untuk menentukan produk laris dan tidak laris pada menu *main course* di *Pizza Hut Delivery* Muara Karang Raya.

f. Pengolahan data metode algoritma C4.5

Pada langkah penelitian pengolahan data, penulis akan melakukan pengolahan data penjualan pertahun dari menu *main course* pizza di *Pizza Hut Delivery* Muara Karang Raya dari proses ini dilakukan menjadikan sebuah informasi. Kumpulan data yang awalnya tidak memiliki sebuah informasi yang dapat disimpulkan jika dilakukan proses penelitian pengolahan data

maka akan menghasilkan sebuah informasi. Setelah dilakukan banyak tahapan diatas maka akan dilanjutkan ke metode algoritma C4.5 untuk melakukan sebuah perhitungan dan menganalisa data yang telah diberikan.

g. Kesimpulan

Berdasarkan tahapan awal sampai akhir dalam data penjualan laris dan tidak laris pada menu *main course* pizza di *Pizza Hut Delivery* Muara Karang Raya menggunakan metode Algoritma C4.5 ini, maka dapat menentukan sebuah kesimpulan sebagai *point – point* penting dalam penelitian sebagai hasil penelitian dari penjualan laris dan tidak laris *main course* memanfaatkan metode algoritma C4.5 untuk memaksimalkan data.

4. HASIL DAN PEMBAHASAN

Hasil pada penelitian menggunakan algoritma C4.5 pada *Pizza Hut Delivery* muara karang raya akan melakukan beberapa langkah – langkah sebagai berikut:

4.1. Data selection

Data yang ada pada database sering kali tidak semuanya dipakai, oleh sebab itu hanya beberapa data saja yang sesuai untuk dianalisis lebih lanjut yang akan diambil dari database[9].

Tabel 1. Tabel data penjualan main course tahun 2021-2022

No	Penjualan	Kode Menu	Nama	Harga	Jumlah Terjual	Status Penjualan
1	Nov 2021	10018	Meat lovers	102.000	120	LARIS
2		10017	Super supreme	102.000	89	TIDAK LARIS
3		10019	American favorite	102.000	106	LARIS
4		10057	Cheese lover	102.000	95	TIDAK LARIS
5		10006	Cheese galore	102.000	171	LARIS
6		10044	Pepperoni	102.000	102	LARIS
7		10027	Chicken lovers	102.000	105	LARIS
8		10041	Veggie garden	102.000	8	TIDAK LARIS
9		10002	Tuna Melt	102.000	121	LARIS
10		10015	Meaty	102.000	106	LARIS
11		10059	Creamy chicken	90.000	101	LARIS
12		10016	Supreme	90.000	104	LARIS
....						
168	Des 2022	10016	Supreme	Rp 90.000	66	TIDAK LARIS

Penulis menggunakan data penjualan main course pizza tahun 2021 – 2022 untuk bertujuan mencari produk dari main course yang mana dari produk tersebut merupakan produk laris dan tidak laris. Selanjutnya dari data yang sudah dipilih tersebut akan disimpan kedalam file tersendiri dalam hal ini penulis menggunakan penyimpanan berupa data file excel.

4.2. Pembagian data *training* dan data *testing*

Berikut ini penulis akan memberikan pembagian data *testing* dan data *training* pada data set yang digunakan untuk penelitian ini:

Gambar 3. Pembagian data *testing* dan data *training*

Berdasarkan pada gambar diatas ini merupakan hasil tampilan dari rasio melalui aplikasi *rapid miner* dimana terdapat didalam sebuah *rasio split data* berupa data *training* dan data *testing* yang dimana sebuah kolom pertama merupakan data *training* dengan jumlah 0,8 atau sebanyak 134 data. Dan kolom kedua adalah data *testing* dengan jumlah 0,2 atau sebanyak 34 data. Maka jumlah keseluruhan data pada pembagian data tersebut sebesar 168 data.

4.3. Transformation data

Transformation atribut akan dilakukan setelah data yang sudah dipilih. Data awal untuk penelitian ini berisi atribut biner yaitu (Tanggal penjualan, kode menu, nama menu, harga, jumlah terjual dan status penjualan), atribut numerik yaitu (Kode menu, harga, dan jumlah terjual). Selanjutnya atribut data yang sudah dipilih akan di transformasi ke dalam bentuk kordinal agar memudahkan tahapan pada proses pemodelan. Berikut ini adalah penjelasan proses transformasi atribut:

Tabel 2. Proses *transformation* data

Nama menu	Harga	Range harga	Jumlah terjual	Range jumlah	Status
A	Murah	$X < 100.000$	Sedikit	$X < 100$	Tidak Laris
B	Mahal	$X \geq 100.000$	Banyak	$X \geq 100$	Laris

Dari tabel 2 penulis melakukan sebuah data *transformation* dengan hasilnya sebagai berikut:

Tabel 3. Data *transformation*

No	Penjualan	Kode Menu	Nama	Harga	Jumlah Terjual	Status Penjualan
1	Nov 2021	10018	Meat lovers	MAHAL	SEDIKIT	LARIS
2		10017	Super supreme	MAHAL	SEDIKIT	TIDAK LARIS
3		10019	American favorite	MAHAL	BANYAK	LARIS
4		10057	Cheese lover	MAHAL	SEDIKIT	TIDAK LARIS
5		10006	Cheese galore	MAHAL	BANYAK	LARIS
6		10044	Pepperoni	MAHAL	BANYAK	LARIS
7		10027	Chicken lovers	MAHAL	BANYAK	LARIS
8		10041	Veggie garden	MAHAL	SEDIKIT	TIDAK LARIS
9		10002	Tuna Melt	MAHAL	BANYAK	LARIS
10		10015	Meaty	MAHAL	BANYAK	LARIS
11		10059	Creamy chicken	MURAH	BANYAK	LARIS
12		10016	Supreme	MURAH	BANYAK	LARIS
....						
168	Des 2022	10016	Supreme	MURAH	SEDIKIT	TIDAK LARIS

Dapat dijelaskan pada tabel 3 adalah sebuah penjelasan dimana nanti terjadinya transformasi data berupa perubahan data pada data numerik menjadi biner dimana atribut HARGA(MURAH dan MAHAL), JUMLAH TERJUAL(SEDIKIT dan BANYAK), dan STATUS PENJUALAN(LARIS dan TIDAK LARIS).

4.4. Perhitungan *entropy* and *gain* dengan *microsoft excel*

Untuk menentukan atribut sebagai akar, maka harus dilakukan perhitungan nilai *entropy* nya terlebih dahulu kemudian selanjutnya menghitung nilai *gain* untuk setiap atribut lalu dipilih nilai *gain* tertinggi. Nilai *gain* yang tertinggi akan dijadikan sebagai akar pohon keputusan [10]. Berikut ini adalah hasil perhitungan *entropy* and *gain* yang akan dijelaskan secara rinci sesuai rumus yang sudah ditetapkan:

Tabel 4. Hasil perhitungan *entropy* and *gain* node1

N1		JK (S)	L (Si)	TL (Si)	Entropy	Gain
Total		168	105	63	0,954434003	
Nama Menu						0,169108
	Meat Lovers	14	14	0		

N1		JK (S)	L (Si)	TL (Si)	Entropy	Gain
	Super Supreme	14	10	4	0,86312056	
	American Favorite	14	11	3	0,749595257	
	Cheese Lover	14	7	7	1	
	Pepperoni	14	11	3	0,749595257	
	Chicken Lovers	14	7	7	1	
	Veggie Garden	14	2	12	0,591672779	
	Tuna Melt	14	10	4	0,863120569	
	Meaty	14	10	4	0,863120569	
	Creamy Chicken	14	4	10	0,863120569	
	Supreme	14	9	5	0,940285959	
Harga						0,021166
	Mahal	140	93	47	0,920658407	
	Murah	28	13	15	0,99631652	
Jumlah Terjual						0,625211
	Banyak	94	93	1	0,08499556	
	Sedikit	74	12	62	0,639457132	

Berdasarkan pada tabel di atas ada sebuah tampilan hasil perhitungan menggunakan microsoft excel dengan metode Algoritma C4.5 yang dimana dari hasil perhitungan tersebut kita bisa lihat bahwa gain tertinggi pada atribut JUMLAH TERJUAL dengan nilai 0,625211 maka atribut JUMLAH TERJUAL menjadi *node* akar dimana terdiri dari sub atribut yaitu BANYAK, dan SEDIKIT. Berdasarkan nilai *entropy* dari ketiga atribut diatas, maka diperlukan dilakukan perhitungan lanjut untuk menentukan *node* akar dari hasil diatas dapat dilihat bahwa faktor dengan nilai *gain* tertinggi. maka dapat dilihat pohon keputusan yang dihasilkan dengan akar semua *node* pada gambar dibawah ini:



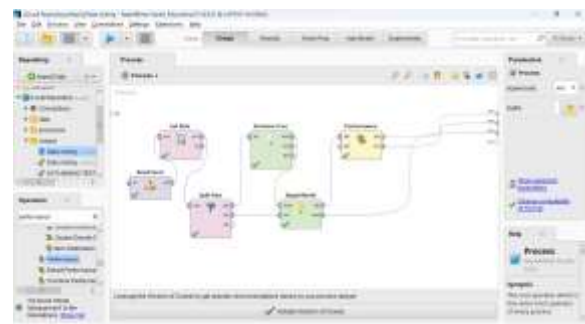
Gambar 4. Pohon keputusan perhitungan manual excel

Pada gambar diatas dapat dijelaskan bahwa setelah melakukan semua perhitungan *entropy* and *gain* yang akan menghasilkan sebuah pohon keputusan seperti gambar diatas yang akan dijadikan acuan pada hasil pohon keputusan yang dijalankan pada aplikasi *RapidMiner* yang bermaksud untuk membuktikan bahwa hasil pohon keputusan perhitungan manual sama atau tidak dengan hasil pohon keputusan yang diimplementasikan di aplikasi *RapidMiner*. Jika kedua hasilnya sama berarti perhitungannya sudah benar.

4.5. Pengujian Algoritma C4.5 dengan aplikasi *Rapid Miner*

Berdasarkan tahapan pada pengujian dengan aplikasi ini, metode ini pada dasarnya merupakan data mining yang akan dilakukan pengujian dengan aplikasi

rapid miner untuk mendapatkan sebuah ilmu pengetahuan yang tersembunyi dan berharga dari data. Metode yang akan digunakan untuk pengujian dengan *rapid miner* adalah algoritma c4.5. Berikut ini adalah pengujian yang akan dilakukan.

Gambar 5. Tampilan pengujian *Rapid Miner*

Pada gambar diatas merupakan langkah – langkah yang dipakai pada *Operators* di *Rapid Miner* untuk mendapatkan sebuah pohon keputusan dan sebuah perhitungan *performance* untuk mengetahui hasil *accuracy* yang didapatkan pada penelitian ini. Berikut ini akan ditampilkan hasil pohon keputusan dan hasil *accuracy* yang didapatkan pada penelitian ini:

Gambar 6. Hasil pohon keputusan *Rapid Miner*

Pada gambar 6 bisa dilihat merupakan hasil dari *decision tree* yang dihasilkan adalah sama dengan hasil perhitungan yang sudah dilakukan secara manual. Maka dari itu bahwa perhitungan yang

dilakukan secara manual sudah dilakukan dengan benar. Setelah mendapatkan sebuah pohon keputusan, maka selanjutnya akan mencari nilai *accuracy* yang dapat ditampilkan sebagai berikut:



Gambar 7. Tampilan nilai *accuracy*

Dari gambar diatas didapatkan sebuah hasil dari RapidMiner dengan nilai *accuracy* sebesar 88,24% dengan data set yang dipakai. Dibawah ini merupakan perhitungan *accuracy* menggunakan *Confusion Matrix* dari gambar diatas. Berikut ini adalah perhitungan yang dilakukan sesuai dengan rumus yang sudah ditentukan:

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

$$Accuracy = \frac{19 + 11}{19 + 2 + 2 + 11} = \frac{30}{34} = 0,8824 = \mathbf{88,24\%}$$

5. KESIMPULAN DAN SARAN

Berlandaskan hasil dari analisa pada penelitian yang sudah dilakukan dengan algoritma C4.5 dapat dibuat sebuah kesimpulan sebagai berikut: Berdasarkan analisa proses bisnis untuk menentukan produk laris dan tidak laris pada *main course* pizza ini telah dinyatakan berhasil setelah dilakukan implementasi dengan menggunakan aplikasi *RapidMiner* serta algoritma C4.5. Faktor yang paling mempengaruhi algoritma C4.5 dalam menentukan produk laris dan tidak laris adalah jumlah terjual yang terjadi setiap transaksi yang terdapat pada data set yang telah dikumpulkan. Hasil yang didapat dari data set yang telah dikumpulkan pada penelitian ini mendapatkan sebuah nilai akurasi 88,24% yang telah dilakukan dengan cara perhitungan dengan aplikasi *RapidMiner* dan juga hitungan manual dan penelitian diatas dapat disimpulkan pula penjualan *main course* laris pada *Pizza Hut Delivery* Muara karang raya adalah menu *American Favorite*, *Meat Lovers*, dan *Pepperonni*.

Dalam melakukan penelitian ini penulis akan memberikan sebuah saran terkait dengan penelitian yang sudah dilakukan yaitu: Penelitian selanjutnya diharapkan akan lebih baik jika menggunakan data terupdate dan juga atribut yang lebih banyak agar dapat menghasilkan sebuah *rule* dan mendapatkan sebuah data yang lebih akurat. Penelitian ini agar bisa di evaluasi lagi kembali dengan metode selain

algoritma C4.5 seperti *Naïve Bayes*, KNN, dan lain – lain agar dibuat sebagai bahan perbandingan.

DAFTAR PUSTAKA

- [1] E. S. Asep Muhidin, “KLASIFIKASI PENJUALAN MAKANAN CEPATSAJI MENGGUNAKAN METODE ALGORITMA C4.5 (STUDIKASUS: AYAM PENYET NABILA),” *J. Teknol. Pelita Bangsa*, vol. 11, no. 1, p. 275, 2020, doi: 10.35393/1730-006-002-014.
- [2] M. Firmansyah and R. Aufany, “Implementasi Metode Decision Tree Dan Algoritma C4.5 Untuk Klasifikasi Data Nasabah Bank,” *Infokam*, vol. XII, no. 1, pp. 1–12, 2016.
- [3] D. Laila Sari, M. Saputra, and H. Gemasih, “Penerapan Data Mining Dalam Proses Prediksi Perceraian Menggunakan Algoritma Naive Bayes Di Kabupaten Aceh Tengah,” *J. Tek. Inform. dan Elektro*, vol. 4, no. 1, pp. 23–35, 2022, doi: 10.55542/jurtie.v4i1.112.
- [4] Y. Mardi, “Data Mining: Klasifikasi Menggunakan Algoritma C4.5,” *Edik Inform.*, vol. 2, no. 2, pp. 213–219, 2017, doi: 10.22202/ei.2016.v2i2.1465.
- [5] Suyanto, *Data Mining untuk Klasifikasi dan Klasterisasi Data*. 2017.
- [6] P. Mayadewi and E. Rosely, “Prediksi Nilai Proyek Akhir Mahasiswa Menggunakan Algoritma Klasifikasi Data Mining,” *Semin. Nas. Sist. Inf. Indones.*, no. November, pp. 329–334, 2015.
- [7] A. Nindy Yuliarina, “Comparison of Prediction Analysis of Gofood Service Users Using the Knn & Naive Bayes Algorithm With Rapidminer Software,” *J. Tek. Inform.*, vol. 3, no. 4, pp. 847–856, 2022, [Online]. Available: <https://doi.org/10.20884/1.jutif.2022.3.4.294>
- [8] R. Fitra and I. Rusdi, “Penerapan Metode Algoritma K-Nearest Neighbor Menggunakan Rapidminer Studio pada Klasifikasi Status Sosial Ekonomi Studi Kasus : Kelurahan Kapuk Muara Rt,” pp. 653–660, 2018.
- [9] T. Syahputra, J. Halim, and E. P. Sintho, “Penerapan Data Mining Dalam Menentukan Pilihan Jurusan Bidang Studi SMA Menggunakan Metode,” *Penerapan Data Min. dalam Menentukan Pilihan Jur. di Bid. Stud. SMA menggunakan Metod. Clust. Dengan Tek. Single Link. JURTEKSI*, vol. IV, no. 2, pp. 1–4, 2018.
- [10] A. Hidayatullah, M. Patria, and H. Septanto, “Pelatihan Microsoft Excel Tingkat Mahir untuk Para Alumni Kejuruan Operator Komputer PPKD Jakarta Timur,” *J. Karya untuk Masy.*, vol. 3, no. 2, pp. 127–137, 2022.