

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/326634448>

Metode Decision Tree Algoritma C.45 Dalam Mengklasifikasi Data Penjualan Bisnis Gerai Makanan Cepat Saji

Article · July 2018

DOI: 10.30645/jurasik.v3i0.60

CITATIONS

4

READS

9,607

2 authors:



Eka Pandu Cynthia

State Islamic University of Sultan Syarif Kasim Riau

10 PUBLICATIONS 8 CITATIONS

[SEE PROFILE](#)



Edi Ismanto

Universitas Muhammadiyah Riau

21 PUBLICATIONS 12 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



GIS and Genetic Algorithm [View project](#)

METODE *DECISION TREE* ALGORITMA C.45 DALAM MENGLASIFIKASI DATA PENJUALAN BISNIS GERAJ MAKANAN CEPAT SAJI

Eka Pandu Cynthia¹, Edi Ismanto²

¹UIN Sultan Syarif Kasim Riau, ²Universitas Muhammadiyah Riau

¹Jalan HR.Soebrantas KM.15 Simpang Baru, Panam, Pekanbaru-Riau,

²Jalan Tuanku Tambusai, Pekanbaru-Riau

¹eka.pandu.cynthia@uin-suska.ac.id, ²edi.ismanto@umri.ac.id

Abstract

Advances in technology and information currently produces smart innovations in business, which can be called business intelligence. One that we can use is Data Mining technology in digging useful information from sales company data warehouse. The purpose of this research is to apply data mining decision decision tree algorithm C4.5 on fast food outlets business and expected to provide information in the form of sales information about food menu that most liked by customers and less popular (bestselling and less in demand). The methodology used in classifying the sales of this research uses the steps of Algorithm C.45, The process uses five steps in KDD (Knowledge Discovery in Databases), which perpetuates activities such as pre-processing, transformation, data mining, interpretation and evaluation. In addition to performing calculations manually, this research case is also tested using Rapidminer application. From the results of the experiment to find data from the sales data of fast food outlets using algorithm C4.5 results of entropy and the highest gain is 1.501991 on the Food Menu attributes on manual calculations. When using the Rapidminer application the results of the results tree as shown in Figure 3.2. Price (s) - Sold Out - Food Menu (Bento Rice = Less Selling, Chest = Laris) Weight (weight) each attribute: Price (0.738), Menu Type (0.067), Sold Number (0.156), Sales Status (0.040).

Keywords: *Bussiness Intelligence, C4.5, Clasification, Data Minig, Decision Tree.*

Abstrak

Kemajuan teknologi dan informasi pada saat ini melahirkan inovasi-inovasi cerdas dalam berbisnis, yang dapat kita sebut kecerdasan bisnis atau business intelligence. Salah satu yang dapat kita manfaatkan adalah teknologi Data Mining dalam menggali informasi yang bermanfaat dari gudang data perusahaan penjualan. Tujuan penelitian ini adalah penulis mencoba menerapkan teknik data mining metode decision tree algoritma C4.5 pada bisnis gerai makanan cepat saji dan diharapkan dapat memberikan informasi berupa klasifikasi penjualan menu makanan yang paling digemari pelanggan dan kurang digemari (laris dan kurang laris). Metodologi yang digunakan dalam mengklasifikasi penjualan menu makanan ini menggunakan tahapan Algoritma C.45, Prosesnya menggunakan lima langkah dalam KDD (Knowledge Discovery in Databases), yang mencakup beberapa aktivitas yaitu seleksi, praproses, transformasi, data mining, interpretasi dan evaluasi. Selain melakukan perhitungan secara manual, kasus penelitian ini juga diujikan menggunakan aplikasi Rapidminer. Dari hasil percobaan pencarian pohon hasil keputusan dari data penjualan gerai makanan cepat saji menggunakan algoritma C4.5 dihasilkan nilai entropy dan gain tertinggi yaitu 1,501991 pada atribut-atribut Menu Makanan pada perhitungan manual. Sedangkan menggunakan aplikasi Rapidminer diperoleh hasil pohon keputusan seperti terlihat pada Gambar 3.2. Harga – Jumlah Terjual – Menu Makanan (Rice Bento = Kurang Laris, Dada = Laris) dengan bobot (weight) masing-masing atribut : Harga (0,738), Jenis Menu (0,067), Jumlah Terjual (0,156), Status Penjualan (0,040).

Kata kunci: *Business Intelligence, C4.5, Data Mining, Decision Tree, Klasifikasi.*

1. PENDAHULUAN

Makan merupakan hal mendasar yang dibutuhkan manusia dalam usaha bertahan hidup, sehingga bermunculan beragam bisnis yang menawarkan aneka jenis makanan. Namun dewasa ini kita rasakan bahwa makanan telah menjelma dari sekedar “kebutuhan bertahan hidup” menjadi “gaya hidup”. Hal ini dapat kita lihat dari banyaknya gerai penjual makanan yang sebenarnya menu makanan tersebut biasa-biasa saja lalu oleh pelaku bisnis kuliner disulap menjadi makanan yang “kekinian” dan akhirnya menjadi favorit segala kalangan usia. Bahkan gaya hidup manusia zaman sekarang yang serba menginginkan kepraktisan dan cepat, melahirkan usaha makanan cepat saji atau yang sering kita sebut *fast food*. Kemajuan teknologi dan informasi pada saat ini melahirkan inovasi-inovasi cerdas dalam berbisnis, yang dapat kita sebut kecerdasan bisnis atau *business intelligence*. Salah satu yang dapat kita manfaatkan adalah teknologi *Data Mining* dalam menggali informasi yang bermanfaat dari gudang data perusahaan penjualan.

Pengambil keputusan berusaha untuk memanfaatkan gudang data yang sudah dimiliki untuk menggali informasi yang berguna membantu mengambil kesimpulan, hal ini menjadikan munculnya cabang ilmu baru untuk mengatasi masalah penggalian informasi atau pola yang penting atau menarik dari data dalam jumlah yang besar, yang disebut dengan *data mining*. Penggunaan teknik *data mining* diharapkan dapat memberikan pengetahuan-pengetahuan baru yang sebelumnya belum diketahui [1].

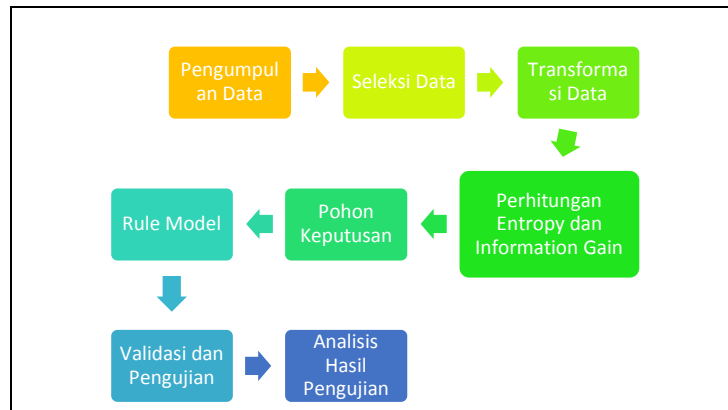
Definisi lain *data mining* adalah serangkaian proses yang memperkerjakan satu atau lebih teknik pembelajaran komputer untuk menganalisis dan mengekstrak pengetahuan secara otomatis atau serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual [2][3]. Penelitian terdahulu tentang pemanfaatan *data mining* pada bisnis diantaranya: pengklasifikasian data nasabah bank untuk pengambilan keputusan pemberian kredit [4][5], prediksi jumlah penjualan motor [6], rekomendasi penerimaan mitra penjualan [7], informasi strategis penjualan batik [5], penjualan produk terlaris distro [8] dan klasifikasi kesetiaan pelanggan terhadap merk produk [9].

Tujuan penelitian ini adalah penulis mencoba menerapkan teknik *data mining* metode *decision tree* algoritma C4.5 pada bisnis gerai makanan cepat saji XYZ dan diharapkan dapat memberikan informasi berupa klasifikasi penjualan menu makanan yang paling digemari pelanggan dan kurang digemari (laris dan tidak laris). Sehingga kedepannya pemilik bisnis ini dapat melakukan analisa menu mengikuti *trend* dan kegemaran pelanggannya.

2. METODOLOGI PENELITIAN

Metodologi yang digunakan oleh penulis dalam mengklasifikasi penjualan menu makanan pada gerai makanan cepat saji XYZ menggunakan tahapan Algoritma C.45, Prosesnya menggunakan lima langkah dalam KDD (*Knowledge Discovery in Databases*), yang mencakup beberapa aktivitas yaitu seleksi, praproses, transformasi, data mining, interpretasi dan evaluasi.

2.1. Tahapan Metodologi Penelitian



Gambar 1. Tahapan Penelitian

A. Pengumpulan Data

Pengumpulan data adalah mengumpulkan data-data yang akan digunakan dalam proses algoritma klasifikasi C4.5.

B. Seleksi Data

Seleksi data adalah memilih data yang akan digunakan dalam proses algoritma klasifikasi C4.5. Tujuan dari seleksi data adalah menciptakan himpunan data target, pemilihan himpunan data, atau memfokuskan pada *subset* variabel atau sampel data, dimana penemuan (*discovery*) akan dilakukan [1][10].

C. Transformasi Data

Transformasi data adalah proses mentransformasi atau mengubah data ke dalam bentuk yang sesuai, agar dapat di proses dengan perhitungan algoritma C4.5..

D. Perhitungan *Entropy* dan *Information Gain*

Perhitungan semua atribut/variabel, *entropy* menggunakan rumus (1) dan *information gain* menggunakan rumus (2) untuk mengetahui *information gain* tertinggi yang akan dijadikan simpul akar pada pembuatan pohon keputusan.

E. Pohon Keputusan (*Decision Tree*)

Pohon keputusan adalah hasil dari proses perhitungan *entropy* dan *information gain*, setelah perhitungan berulang-ulang sampai semua atribut pohon memiliki kelas dan tidak bisa lagi dilakukan proses perhitungan.

F. Aturan-aturan/ *Rule Model*

Aturan-aturan/ *Rule model* adalah uraian penjelasan yang merepresentasikan sebuah pohon keputusan.

G. Validasi dan Pengujian

Validasi dan pengujian adalah Pengujian yang dilakukan untuk mengetahui semua fungsi bekerja dengan baik atau tidak. Validasi dilakukan dengan *Ten-fold Cross Validation*. *Ten-fold Cross Validation* adalah validasi yang dilakukan dengan cara membagi suatu set data menjadi sepuluh segmen yang berukuran sama besar dengan cara melakukan pengacakan data. Validasi dan pengujian dilakukan untuk mengetahui tingkat akurasi, presisi, dan recall dari hasil prediksi klasifikasi. Akurasi adalah persentase dari catatan yang diklasifikasikan dengan benar dalam pengujian dataset. Presisi adalah persentase data yang diklasifikasikan sebagai model baik yang sebenarnya juga baik. *Recall* adalah pengukuran tingkat pengenalan positif sebenarnya.

H. Analisis Hasil Pengujian

Analisa yang dilakukan untuk memastikan bahwa hasil pengujian benar-benar sesuai dengan pembahasan. Analisa dilakukan dengan melakukan perhitungan kembali hasil validasi dan pengujian (akurasi, presisi, dan recall) secara manual. Apakah perhitungan yang dilakukan akan menghasilkan nilai yang sama atau tidak, dibantu dengan *Confusion matrix*. *Confusion Matrix* adalah model yang akan membentuk matrix yang terdiri dari *true positif* atau tupel positif dan *true negatif* atau tupel negatif. *Confusion matrix* berisi informasi aktual (actual) dan prediksi (*predicted*) pada sistem klasifikasi. *Confusion matrix* dapat memvisualisasi kinerja algoritma klasifikasi.

2.2. Analisis Data Mining

Data mining adalah bagian integral dari penemuan pengetahuan dalam *database* yang merupakan proses dengan urutan sebagai berikut [11][12] :

1. *Data cleaning* (untuk menghilangkan noise dan inkonsistensi data)
2. *Data integration* (beberapa *data sources* akan dikombinasikan)
3. *Data selection* (hanya data yang dapat dipakai untuk analisis saja yang akan diambil dari database)
4. *Data transformation* (data akan ditransformasikan ke bentuk yang lebih terstruktur untuk mempermudah proses *data mining*)
5. *Data mining* (proses utama data mining dimana teknik *data mining* diterapkan)
6. *Pattern evaluation*
7. *Knowledge presentation* (dimana visualisasi dan representasi hasil diberikan kepada pengguna)

Tujuan utama penerapan *data mining* adalah untuk prediksi (*prediction*) dan uraian (*description*). Klasifikasi adalah proses menemukan model (atau fungsi) yang akan mengelompokkan kelas data sehingga dapat memprediksi kelas obyek yang tidak diketahui. Model yang diperoleh bias diwakili diberbagai bentuk, seperti klasifikasi berbentuk aturan IF-THEN, pohon keputusan, rumus matematika, atau jaringan saraf tiruan. *Rule based classification* adalah model berbentuk seperangkat

aturan berbentuk IF-THEN. Aturan berbentuk IF-THEN adalah cara yang baik untuk merepresentasikan pengetahuan. Aturan dapat diekstraksi dari pohon keputusan.

2.3. Algoritma C4.5

Algoritma C4.5. merupakan kelompok algoritma pohon keputusan (*decision tree*). Algoritma ini mempunyai input berupa *training samples* dan *samples*. *Training samples* berupa data contoh yang akan digunakan untuk membangun sebuah *tree* yang telah diuji kebenarannya. Sedangkan *samples* merupakan *field-field* data yang nantinya akan kita gunakan sebagai parameter dalam melakukan klasifikasi data [7].

Algoritma C4.5 adalah salah satu metode untuk membuat *decision tree* berdasarkan training data yang telah disediakan. Algoritma C4.5 merupakan pengembangan dari ID3. Beberapa pengembangan yang dilakukan pada C4.5 adalah sebagai antara lain bisa mengatasi *missing value*, bisa mengatasi *continue data*, dan *praining*. Secara garis besar langkah-langkah yang dilakukan oleh Algoritma C.45 dalam membentuk pohon keputusan adalah sebagai berikut :

Input : sample training, label training, atribut

1. Membuat simpul akar untuk pohon yang dibuat
2. Jika semua sampel positif, berhenti dengan suatu pohon dengan satu simpul akar, beri tanda (+)
3. Jika semua sampel negatif, berhenti dengan suatu pohon dengan satu simpul akar, beri tanda (-)
4. Jika atribut kosong, berhenti dengan suatu pohon dengan suatu simpul akar, dengan label sesuai nilai yang terbanyak yang ada pada label training
5. Untuk yang lain, Mulai
 - a. A ---- atribut yang mengklasifikasikan sampel dengan hasil terbaik (berdasarkan *gain rasio*)
 - b. Atribut keputusan untuk simpul akar ---- A
 - c. Untuk setiap nilai, v_i , yang mungkin untuk A
 - 1) Tambahkan cabang di bawah akar yang berhubungan dengan $A = v_i$
 - 2) Tentukan sampel S_{v_i} sebagai *subset* dari sampel yang mempunyai nilai v_i untuk atribut A
 - 3) Jika sampel S_{v_i} kosong
 - i. Di bawah cabang tambahkan simpul daun dengan label = nilai yang terbanyak yang ada pada label training
 - ii. Yang lain tambah cabang baru di bawah cabang yang sekarang C 4.5 (*sample training, label training, atribut – [A]*).
 - d. Berhenti

Output : Decision Tree.

Mengubah *tree* yang dihasilkan dalam beberapa *rule*. Jumlah *rule* sama dengan jumlah *path* yang mungkin dapat dibangun dari *root* sampai *leaf node*. *Tree Praining* dilakukan untuk menyederhanakan *tree* sehingga akurasi dapat bertambah. *Pruning* ada dua pendekatan, yaitu :

- 1) *Pre-praining*, yaitu menghentikan pembangunan suatu *subtree* lebih awal (yaitu dengan memutuskan untuk tidak lebih jauh mempartisi *data training*). Saat seketika berhenti, maka *node* berubah menjadi *leaf* (node akhir). Node akhir ini menjadi kelas yang paling sering muncul di antara *subset* sampel.
- 2) *Post-praining*, yaitu menyederhanakan *tree* dengan cara membuang beberapa cabang *subtree* setelah *tree* selesai dibangun. *Node* yang jarang dipotong akan menjadi *leaf* (node akhir) dengan kelas yang paling sering muncul.

Secara umum algoritma C4.5 adalah untuk membangun pohon keputusan adalah sebagai berikut :

1. Pilih atribut sebagai akar
2. Buat cabang untuk masing-masing nilai
3. Bagi kasus dalam cabang
4. Ulangi proses untuk masing-masing cabang sampai semua kasus pada cabang memiliki kelas yang sama.

2.4. Entropy dan Information Gain

Sebuah obyek yang diklasifikasikan dalam pohon harus dites nilai *Entropy* -nya. *Entropy* adalah ukuran dari teori informasi yang dapat mengetahui karakteristik dari *impurity* dan *homogeneity* dari kumpulan data. Dari nilai *Entropy* tersebut kemudian dihitung nilai *information gain* (IG) masing-masing atribut. *Entropy* (S) merupakan jumlah bit yang diperkirakan dibutuhkan untuk dapat mengekstrak suatu kelas (+ atau -) dari sejumlah data acak pada ruang sampel S. *Entropy* dapat dikatakan sebagai kebutuhan bit untuk menyatakan suatu kelas [8]. Semakin kecil nilai *Entropy* maka akan semakin *Entropy* digunakan dalam mengekstrak suatu kelas. *Entropy* digunakan untuk mengukur ketidakkaslian S.sistem informasi atau disebut dengan *processing system*.

Pemilihan atribut sebagai simpul, baik akar maupun simpul internal didasarkan pada nilai *Gain* tertinggi dari atribut-atribut yang ada. Rumus menghitung nilai *entropy* menggunakan Persamaan 1 [3] :

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

Keterangan dari Persamaan 1 :

S : himpunan kasus.

n : jumlah partisi S

Pi : proporsi Si terhadap S

Information gain adalah salah satu *attribute selection measure* yang digunakan untuk memilih *test attribute* tiap *node* pada *tree*. Atribut dengan informasi *gain* tertinggi dipilih sebagai *test* atribut dari suatu *node*. Gain (S,A) merupakan perolehan informasi dari atribut A *relative* terhadap *output* data S. Perolehan informasi didapat dari *output* data atau *variable dependent* S yang dikelompokkan berdasarkan atribut A, dinotasikan dengan gain (S,A). Menghitung nilai *information gain* menggunakan Persamaan 2 :

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

Keterangan dari Persamaan 2 :

S : himpunan kasus.

A : Atribut.

n : jumlah partisi atribut A.

|S_i| : jumlah kasus pada partisi ke i.

|S| : jumlah kasus dalam S.

3. HASIL DAN PEMBAHASAN

3.1. Analisa Data

Data yang digunakan adalah contoh data penjualan pada sebuah gerai makanan cepat saji dalam satu bulan (30 hari).

Tabel 1. Data Penjualan

No.	Tanggal Penjualan	Kode Menu	Jenis Menu	Nama Menu	Harga	Jumlah Terjual	Status Penjualan
1	1-Jun-2018	SM01	Makanan	Super Mumer	20000	15	Laris
2		PM01	Makanan	Paket Mumer 1	16000	10	Kurang Laris
3		PM02	Makanan	Paket Mumer 2	19000	13	Laris
4		PM03	Makanan	Paket Mumer 3	23000	12	Kurang Laris
5		PM04	Makanan	Paket Mumer 4	27000	10	Kurang Laris
6		RB01	Makanan	Rice Bento	10000	15	Kurang Laris
7		OB01	Makanan	Original Burger	8000	15	Kurang Laris
8		AG01	Makanan	Ayam Geprek	15000	17	Laris
9		FF01	Makanan	French Fries	8000	10	Kurang Laris
10		DD01	Makanan	Dada	10000	25	Kurang Laris
11		PA01	Makanan	Paha Atas	9000	15	Kurang Laris
12		SY01	Makanan	Sayap	8000	19	Kurang Laris
13		PB01	Makanan	Paha Bawah	8000	16	Kurang Laris
14		NP01	Makanan	Nasi Putih	4000	45	Laris
15		IF01	Minuman	Ice Float Mocca	7000	7	Kurang Laris
16		IF02	Minuman	Ice Float Tropical	7000	9	Kurang Laris
17		IF03	Minuman	Ice Float Lechy	7000	11	Kurang Laris
18		IF04	Minuman	Ice Float Orange	7000	15	Kurang Laris
19		IF05	Minuman	Ice Float Mango	7000	13	Kurang Laris
20		SIC01	Minuman	Sundae Ice Cream	5000	25	Laris
21		SD01	Minuman	Pepsi	5000	35	Kurang Laris
22		SD02	Minuman	Mirinda	5000	29	Laris
23		AM01	Minuman	Air Mineral	4000	35	Kurang Laris
:	:	:	:	:	:	:	:
668	30-Jun-2018	SM01	Makanan	Super Mumer	20000	24	Kurang Laris
669		PM01	Makanan	Paket Mumer 1	16000	17	Laris
670		PM02	Makanan	Paket Mumer 2	19000	9	Kurang Laris
671		PM03	Makanan	Paket Mumer 3	23000	11	Laris
672		PM04	Makanan	Paket Mumer 4	27000	29	Kurang Laris
673		RB01	Makanan	Rice Bento	10000	17	Kurang Laris
674		OB01	Makanan	Original Burger	8000	27	Kurang Laris
675		AG01	Makanan	Ayam Geprek	15000	30	Laris

No.	Tanggal Penjualan	Kode Menu	Jenis Menu	Nama Menu	Harga	Jumlah Terjual	Status Penjualan
676		FF01	Makanan	French Fries	8000	10	Kurang Laris
677		DD01	Makanan	Dada	10000	29	Laris
678		PA01	Makanan	Paha Atas	9000	29	Kurang Laris
679		SY01	Makanan	Sayap	8000	17	Laris
680		PB01	Makanan	Paha Bawah	8000	10	Laris
681		NP01	Makanan	Nasi Putih	4000	33	Kurang Laris
682		IF01	Minuman	Ice Float Mocca	7000	7	Kurang Laris
683		IF02	Minuman	Ice Float Tropical	7000	8	Laris
684		IF03	Minuman	Ice Float Lechy	7000	7	Kurang Laris
685		IF04	Minuman	Ice Float Orange	7000	10	Kurang Laris
686		IF05	Minuman	Ice Float Mango	7000	13	Laris
687		SIC01	Minuman	Sundae Ice Cream	5000	33	Kurang Laris
688		SD01	Minuman	Pepsi	5000	22	Laris
689		SD02	Minuman	Mirinda	5000	25	Kurang Laris
690		AM01	Minuman	Air Mineral	4000	30	Laris

Data penjualan tersebut lalu diklasifikasikan menjadi 5 bagian, yaitu pembagian berdasarkan Jenis Menu, Menu Makanan, Menu Minuman, Harga dan Jumlah Terjual.

Tabel 2. Klasifikasi Berdasarkan Jenis Menu

JENIS MENU	STATUS PENJUALAN	
	LARIS	KURANG LARIS
Makanan	165	255
Minuman	111	159

Tabel 4. Klasifikasi Berdasarkan Jumlah Terjual

RENTANG JUMLAH	STATUS PENJUALAN	
	LARIS	KURANG LARIS
<= 25 Pcs	190	334
> 25 Pcs	86	80

Tabel 3. Klasifikasi Berdasarkan Menu Minuman

Menu Minuman	STATUS PENJUALAN	
	LARIS	KURANG LARIS
Ice Float Mocca	11	19
Ice Float Tropical	12	18
Ice Float Lechy	12	18
Ice Float Orange	15	15
Ice Float Mango	8	22
Sundae Ice Cream	9	21
Pepsi	15	15
Mirinda	11	19
Air Mineral	18	12

Tabel 5. Klasifikasi Berdasarkan Harga

RENTANG HARGA	STATUS PENJUALAN	
	LARIS	KURANG LARIS
<= 25.000	263	397
> 25.000	13	17

Tabel 6. Klasifikasi Berdasarkan Menu Makanan

Menu Makanan	STATUS PENJUALAN	
	LARIS	KURANG LARIS
Super Mumer	13	17
Paket Mumer 1	11	19
Paket Mumer 2	11	19
Paket Mumer 3	12	18
Paket Mumer 4	13	17
Rice Bento	15	15
Original Burger	6	24

Menu Makanan	STATUS PENJUALAN	
	LARIS	KURANG LARIS
Ayam Geprek	13	17
French Fries	10	20
Dada	14	16
Paha Atas	10	20

Menu Makanan	STATUS PENJUALAN	
	LARIS	KURANG LARIS
Sayap	15	15
Paha Bawah	8	22
Nasi Putih	14	16

3.2. Proses Data Mining

3.2.1. Perhitungan Algoritma C4.5

Perancangan *data mining* pada penelitian ini menggunakan algoritma C4.5. Proses pada pohon keputusan adalah mengubah bentuk data (*table*) menjadi model pohon, mengubah model pohon menjadi *rule*, dan menyederhanakan *rule*. Kasus yang ada pada Tabel 7 Data Penjualan akan dibuat pohon keputusan untuk menentukan menu makanan yang terjual dengan status penjualannya kurang laris, laris dan sangat laris berdasarkan jenis menu dan nama-nama menu yang ada. Berikut ini adalah penjelasan lebih terperinci mengenai tiap-tiap langkah dalam pembentukan pohon keputusan dengan menggunakan algoritma C4.5:

1. Menghitung : jumlah kasus untuk keputusan Kurang Laris, jumlah kasus untuk keputusan Laris, jumlah kasus untuk keputusan Sangat Laris, menghitung *entropy* untuk semua kasus, menghitung *entropy* untuk sub kasus (berdasarkan atribut) dan menghitung *gain* untuk setiap atribut. Hasil perhitungan untuk node awal dapat dilihat pada table 3.5 di bawah ini.

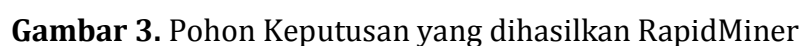
Tabel 7. Perhitungan Node Awal

Node		Jumlah Kasus	Status Penjualan		Nilai	
			Laris	Kurang Laris	Entropy	Gain
		(S)	(S1)	(S2)		
1	Total	690	276	414	0.970951	
	Jenis Menu					0.76491
	Makanan	420	165	255	0.966619	
	Minuman	270	111	159	0.977080	
	Menu Makanan					1.501991
	Super Mumer (A)	30	13	17	0.987138	
	Paket Mumer 1 (B)	30	11	19	0.948078	
	Paket Mumer 2 (C)	30	11	19	0.948078	
	Paket Mumer 3 (D)	30	12	18	0.970951	
	Paket Mumer 4 (E)	30	13	17	0.987138	
	Rice Bento (F)	30	15	15	1	
	Original Burger (G)	30	6	24	0.721928	
	Ayam Geprek (H)	30	13	17	0.987138	
	French Fries (I)	30	10	20	0.918296	
	Dada (J)	30	14	16	0.996792	
	Paha Atas (K)	30	10	20	0.918296	
	Sayap (L)	30	15	15	1	
	Paha Bawah (M)	30	8	22	0.836641	
	Nasi Putih (N)	30	14	16	0.996792	

2. Pohon keputusan yang terbentuk pada perhitungan manual ditunjukkan pada Gambar 2 berikut :



Pengujian menggunakan Rapid Miner menghasilkan pohon keputusan seperti di bawah ini :



4. SIMPULAN

Dari hasil percobaan pencarian pohon hasil keputusan dari data penjualan gerai makanan cepat saji menggunakan algoritma C4.5 dihasilkan nilai *entropy* dan *gain* tertinggi yaitu 1,501991 pada *atribut-atribut* Menu Makanan pada perhitungan manual. Sedangkan menggunakan aplikasi Rapidminer diperoleh hasil pohon keputusan seperti terlihat pada Gambar 3.2. Harga – Jumlah Terjual – Menu Makanan (Rice Bento = Kurang Laris, Dada = Laris) dengan bobot (*weight*) masing-masing atribut : Harga (0,738), Jenis Menu (0,067), Jumlah Terjual (0,156), Status Penjualan (0,040).

DAFTAR PUSTAKA

- [1]. Arifin, M.F. dan Firianah, Devi., “Penerapan Algoritma Klasifikasi C4.5 dalam Rekomendasi Penerimaan Mitra Penjualan Studi Kasus : PT Atria Artha Persada” IncomTech, Jurnal Telekomunikasi dan Komputer, vol.8, no.2 Halaman 87-102, ISSN : 2085-4811, 2018.
- [2]. Ariwibowo, A.S., “Metode Data Mining Untuk Klasifikasi Kesetiaan Pelanggan Terhadap Merek Produk” Seminar Nasional Sistem Informasi Indonesia, Halaman 532-540, 2-4 Desember 2013.
- [3]. Azwanti, Nurul., “Analisa Algoritma C4.5 Untuk Memprediksi Penjualan Motor Pada PT. Capella Dinamik Nusantara Cabang Muka Kuning”, Jurnal Ilmiah Ilmu Komputer, Informatika Mulawarman Vol 13 No. 1 Halaman 33-38, E-ISSN : 2597-4963 P-ISSN : 1858-4853, Februari 2018.
- [4]. Hiriana, Nadiya dan Rasyidan, Muhammad., “Penerapan Metode Decision Tree Algoritma C4.5 untuk Seleksi Calon Penerima Beasiswa Tingkat Universitas” Al Ulum Sains dan Teknologi, Vol. 3 No. 1 Halaman 9-13, November 2017.
- [5]. Nugroho, Y.S. dan Alirsyadi, F.Y., “Implementasi Data Mining Sebagai Informasi Strategis Penjualan Batik (Studi Kasus Batik Mahkota Laweyan)” Prosiding SNATIF ke-2, Halaman 161-168, ISBN: 978-602-1180-21-1, Universitas Muria Kudus, 2015.
- [6]. Pramono, Fajar., et al., “Komparasi Klasifikasi Penentuan Keterlambatan Siswa SMA Datang Upacara Menggunakan Algoritma C4.5” Seminar Nasional Teknologi Informasi dan Komunikasi (Sentika), Halaman 80-86, ISSN : 2089-9815, Yogyakarta 23-24 Maret 2018.
- [7]. Rahman, Arif Muhammad., “Algoritma C4.5 Untuk Menentukan Mahasiswa Penerima Beasiswa (Studi Kasus PPS IAIN Raden Intan Bandar Lampung)” Jurnal Tim Darmajaya Vol. 1 No. 2 Halaman 118-128, ISSN : 2442-5567, Oktober 2015.
- [8]. Rani, L.N., “Klasifikasi Nasabah Menggunakan Algoritma C4.5 Sebagai Dasar Pemberian Kredit” Jurnal Inovtek Polbeng – Seri Informatika Vol. 1 No. 2 Halaman 126-132, ISSN : 2527-9866, November 2016.
- [9]. Rusito dan T.M. Meidy, “Implementasi Metode Decision Tree dan Algoritma C4.5 Untuk Klasifikasi Data Nasabah Bank” Infokam Nomor I Th. XII Halaman 1-12, Maret 2016.
- [10]. Susanto, et al., “Penerapan Data Mining Classification Untuk Prediksi Perilaku Pola Pembelian Terhadap Waktu Transaksi Menggunakan Metode Naïve Bayes” Konferensi Nasional Sistem dan Informatika STMIK STIKOM, Halaman 313-318, Bali, 9-10 Oktober 2015.
- [11]. Tyas, A.F. et al., “Klasifikasi Data Dengan Menggunakan Algoritma C4.5 dan TAN” Tugas Akhir Program Studi S1 Teknik Informatika, Fakultas Teknik Informatika, Universitas Telkom, Halaman 15-61, 2010.

- [12]. Winata, A.D., "Analisis dan Prediksi Penjualan Produk Terlaris Distro RootShoes Dengan Aplikasi Android" Artikel Publikasi Universitas Muhammadiyah Surakarta, 2017.