

Penerapan Data Mining Untuk Memprediksi Ketersediaan Stok Produk HNI HPAI Menggunakan Algoritma C4.5

Murlena¹, Diwi Apriana²

¹²Fakultas Teknik, Program Studi Ilmu Komputer

¹²Universitas Pat Petulai, Bengkulu, Indonesia

Email : ana.murlena@gmail.com, dapriana102@gmail.com

Article Information

Article history

Received 8 April 2022

Revised 15 May 2022

Accepted 1 June 2022

Available 25 June 2022

Keywords

Data Mining
Algoritma C4.5
Stock
Hpai

Corresponding Author:

Murlena,
Prodi Ilmu Komputer,
Universitas Pat Petulai,
Email : ana.murlena@gmail.com

ABSTRACT

In a company, the availability of this stock must be considered to avoid product losses caused by expiration and also to avoid order cancellations. The data contained in the HPAI company application, called the HSIS application, can be utilized and applied for data mining in determining which products are in demand in everyday life. To maintain the availability of product stock, the purpose of this study is to test the Stock Agent using the C4.5 Algorithm to predict that the availability of products that must be maintained is in accordance with customer needs. The research method used in this study uses Data Mining techniques with analysis using the C4.5 algorithm. It is hoped that the use of Analysis using the C4.5 Algorithm on Stock Agents can overcome stockpiling and finance or Stock Agent capital can be used to buy stock that must be available. The results of the analysis using the C4.5 algorithm state that even though the product is expensive, the best-selling product is herbs product. So of the three types of HPAI HNI products (Herbs product, Health food & beverage and Cosmetic & home care), the stock that must be increased or available is the Herbs Product.

Keywords : *data mining, algoritma C4.5, stock, hpai*

ABSTRAK

Pada suatu perusahaan, ketersediaan stok ini harus diperhatikan untuk menghindari kerugian produk yang disebabkan karena *expired* dan juga untuk menghindari terjadinya pembatalan orderan. Data yang terdapat pada aplikasi perusahaan HPAI yang dinamakan aplikasi HSIS ini dapat dimanfaatkan dan diterapkannya *data mining* dalam menentukan produk-produk apa saja yang laris dalam kehidupan sehari-hari. Untuk menjaga ketersediaan stok produk, maka tujuan penelitian ini adalah melakukan pengujian pada AgenStok menggunakan Algoritma C4.5 untuk memprediksi agar produk yang harus dijaga ketersediaannya sesuai dengan kebutuhan *customer*. Metode penelitian yang digunakan dalam penelitian ini menggunakan teknik *Data Mining* dengan analisis menggunakan algoritma C4.5. Diharapkan penggunaan Analisis menggunakan Algoritma C4.5 pada Agenstok bisa mengatasi penumpukan stok dan keuangan atau modal Agenstok dapat dimanfaatkan untuk membeli stok yang harus tersedia. Hasil dari analisa menggunakan algoritma C4.5 menyatakan bahwa meskipun harga produknya mahal namun tetap laris adalah *herbs product* (produk herbal). Jadi dari ketiga jenis produk HNI HPAI (*Herbs product, Health food & beverage dan Cosmetic & home care*) maka yang stoknya harus dibanyakkan atau tersedia adalah *Herbs Product*.

Kata Kunci : *data mining, algoritma C4.5, stok, hpai*

Copyright@2022 Murlena, Diwi Apriana

This is an open access article under the [CC-BY-NC-SA](https://creativecommons.org/licenses/by-nc-sa/4.0/) license.



1. Pendahuluan

Perkembangan teknologi dan komunikasi saat ini menuntut semua aspek untuk beradaptasi dalam semua bidang seperti bidang pendidikan, industri, ekonomi serta perdagangan (Widyastuti et al. 2021). Produk HPAI ini merupakan produk herbal yang saat ini laris dipasarkan karena produk-produknya diakui bagus dan berkualitas. Untuk lebih meningkatnya kinerja pada perusahaan dibidang pelayanan dan kemudahan dalam mengakses informasi dan juga transaksi yang *real time* bagi agen dan *stakeholder* maka HNI-HPAI menggunakan beberapa teknologi, salah satunya *HSIS (HNI-HPAI Sales Integrated System)*. *HSIS* ini mengintegrasikan transaksi yang dilakukan secara *online* dengan beberapa fitur dan informasi yang mudah diakses secara *real time* diantaranya mengenai peningkatan *omzet*, berapa banyak ketersediaan saldo produk dan juga bagaimana perkembangan jumlah agen per hari.

Dikarenakan produk-produk herbal ini ada batas penggunaannya maka sebelum *expired* produk tersebut harus habis. Untuk mengatasi agar tidak terjadinya penumpukan produk pada agenstok maka dilakukan penerapan *data mining* untuk memprediksi ketersediaan stok. Dengan tidak menumpuknya produk pada Agenstok maka perputaran keuangan atau modal pun dapat teratasi (Fikri and Verina 2020). Maka dengan menggunakan Algoritma C4.5 ini dapat diprediksi produk apa saja yang harus disediakan lebih banyak dan produk apa yang tidak mesti disediakan, sehingga dengan stok yang sesuai dapat memenuhi kebutuhan konsumen dan agenstok pun terhindar dari pembatalan orderan yang dapat merugikan agenstok itu sendiri (Jumiyati and Prasetyaningrum 2020).

2. Kajian Terdahulu

2.1 Landasan Teori

Data Mining

Secara umum data mining terdiri atas 2 kata yaitu : data dan mining. Data merupakan kumpulan fakta yang bisa berupa objek atau kejadian yang tercatat dan terekam yang belum mempunyai arti, sedangkan kata mining merupakan proses penambangan. Dengan demikian data mining dapat didefinisikan sebagai suatu proses penambangan data yang menghasilkan *output* berupa ilmu pengetahuan. Data mining merupakan bagian dari beberapa bidang ilmu pengetahuan seperti *machine learning*, pengenalan pola, statistik, basis data dan juga visualisasi dalam penanganan masalah untuk pengambilan informasi dari *basis data* yang besar (larose, 2005).

Algoritma C4.5

Dalam proses pemecahan masalah dan penemuan pengetahuan baru ada beberapa klasifikasi yang dilakuakn secara umum diantaranya: Estimasi, Association,

Klasifikasi, clustering dan Prediksi. Terdapat beberapa teknik klasifikasi yang dipergunakan sebagai solusi pemecahan masalah diantaranya : Algoritma C4.5, Algoritma *K-Nearest Neighbor*, ID3, *Naïve Bayesian Clasification*, CART (*Clasificication And Regression Tree*). Algoritma C4.5 adalah algoritma yang dipergunakan untuk membentuk pohon keputusan. Pada algoritma C4.5 untuk menyelesaikan masalah ada beberapa elemen yang diketahui yaitu : *Entropy* dan *Gain*. *Entropy* merupakan perkiraan jumlah bit yang dibutuhkan untuk bisa mengekstrak suatu kelas (negatif atau positif) dari sejumlah data acak pada ruang sampel *S*. *Gain* merupakan informasi yang diperoleh dari atribut *A* relative terhadap keluaran data *S*.

Stok

Stok merupakan barang dagang yang disimpan untuk kemudian dijual dalam operasi bisnis perusahaan. Dan bahan yang digunakan dalam proses produksi atau yang disimpan untuk tujuan itu (warren,2005). Stok adalah bahan atau barang yang disimpan yang akan digunakan untuk memenuhi tujuan tertentu (Eddy herjanto, 2007). Stok dalam kehidupan sehari-hari misalnya untuk digunakan dalam proses produksi atau perakitan, untuk dijual kembali, atau untuk suku cadang dari suatu peralatan atau mesin.

HPAI

Hpai ini singkatan dari Herba Penawar Al wahida Indonesia. Hpai merupakan salah satu dari perusahaan bisnis Halal Network yang ada di Indonesia yang memfokuskan produknya pada produk-produk herbal. PT HPAI mempunyai tujuan menjayakan produk-produk yang halal dan juga berkualitas yang berazaskan Thibunnabawi. HPAI ini juga mempunyai misi membumikan dan memajukan serta mengaktualisasikan perekonomian Islam di negara Indonesia melalui enterpreneurship.

2.2 Penelitian Terdahulu

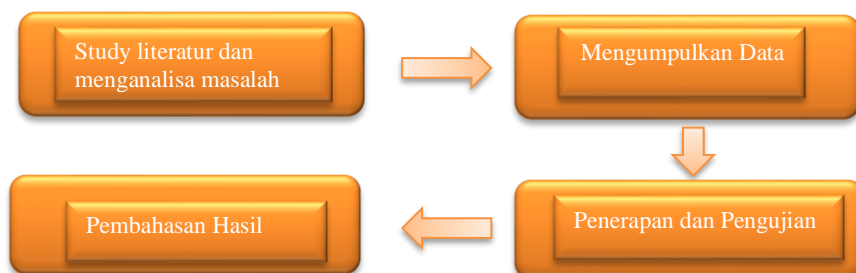
Penelitian yang dilakukan oleh Gunartatik Esthining tyas dkk, yang berjudul Penerapan Algoritma C4.5 Untuk Menentukan Persediaan Obat (Studi Kasus Di RS Bethesda Yogyakarta) bahwa penelitian ini menggambarkan “persediaan obat adalah salah satu faktor penting. Jumlah persediaan atau stok harus dipertimbangkan, dengan tujuan untuk menjaga stok tersedia stabil. Stok tersedia berarti bahwa di *warehouse* tidak ada kelebihan stok atau kekurangan stok. Dalam penelitian ini yang dimaksud adalah *warehouse* Farmasi Rawat Jalan Lantai I Rumah Sakit Bethesda. Permasalahannya adalah bahwa setiap item obat memiliki pengeluaran yang berbeda, sangat tergantung pada peresepan dokter. Maka perhitungan diperlukan untuk menentukan pola obat apa yang perlu disediakan (*reorder*), yang tidak perlu disediakan atau bahkan perlu dikurangi jumlahnya di *warehouse*. Penerapan algoritma C4.5 dapat digunakan untuk memprediksi

ketersediaan stok di Farmasi Rawat Jalan Bawah Lantai 1 Rumah Sakit Bethesda. Hasil penghitungan nilai *entropy* dalam algoritma C4.5 menggunakan *Decision tree RapidMiner* dapat menghasilkan pola apakah stok di Farmasi Rawat Jalan Bawah Lantai 1 Rumah Sakit Bethesda perlu ditambah, cukup atau perlu dikurangi (Esthiningtyas, 2020).

Penelitian yang dilakukan oleh Nova Hayati dkk, yang berjudul *Optimalisasi Prediksi Penjualan Produk Herbal Menggunakan Metode Monte Carlo dalam Meningkatkan Transaksi (Studi Kasus: Toko Herbal An Nabawi)* bahwa dalam penelitiannya menggambarkan bahwa “banyaknya peminat produk herbal maka diperlukan ketersediaan produk untuk meningkatkan transaksi penjualan produk tersebut. Untuk meningkatkan transaksi penjualan produk tersebut maka salah satu proses yang dapat dilakukan yaitu memprediksi penjualan produk herbal, dengan data yang digunakan dari Januari 2018 sampai Desember 2019 pada toko herbal An Nabawi. Proses prediksi dilakukan dengan metode Monte Carlo dan untuk mempermudah proses prediksi diimplementasikan sebuah sistem berbasis web dengan bahasa pemrograman PHP. Dengan metode Monte Carlo yang digunakan pada penelitian ini untuk memprediksi penjualan produk herbal sehingga dapat dimanfaatkan oleh pihak pimpinan untuk melakukan keputusan terhadap ketersediaan produk herbal di toko tersebut. Hasil prediksi penjualan yang didapatkan dari proses simulasi Monte Carlo dengan tingkat akurasi sebesar 87,91%. Dengan hal itu metode Monte Carlo dapat diterapkan untuk memprediksi penjualan produk herbal pada masa yang akan datang dan dapat dimanfaatkan oleh pimpinan toko untuk mengambil keputusan dalam ketersediaan produk herbal (Hayati, N 2020).

3. Metodologi Penelitian

Metodologi penelitian menjadi dasar peneliti dalam melakukan penelitian, maka sesuai dengan tujuan penelitian dan mempelajari beberapa literatur maka penulis memutuskan metode yang cocok digunakan untuk memecahkan masalah ini menggunakan teknik *data mining* dengan algoritma C4.5. Supaya penelitian ini lebih terarah dan sistematis maka kerangka penelitian ini dapat dilihat pada gambar 1 dibawah ini.



Gambar 1. Kerangka Penelitian

Sesuai kerangka penelitian pada gambar 1 di atas maka dapat diuraikan sebagai berikut:

a. Studi Literatur dan Menganalisa Masalah

Pada penelitian ini pengujian dilakukan pada data member Agen stok HPAI untuk menentukan ketersediaan stok produk HPAI pada agen stok HPAI.

b. Mengumpulkan Data

Langkah dalam mendapatkan data-data yang diperlukan dalam rangka mencapai tujuan penelitian. Pengumpulan data penjualan yang telah dilakukan dalam beberapa bulan terakhir yang berasal dari agen stok HPAI pada aplikasi HSIS.

c. Penerapan dan Pengujian

Pada langkah penerapan dan pengujian ini merupakan proses eksekusi data penjualan agen stok HPAI dengan menganalisa menggunakan Algoritma C4.5 sehingga nantinya akan didapatkan hasil pengujian berupa hasil prediksi ketersediaan stok produk HPAI pada agen stok sehingga para agen stok dapat memutuskan kebutuhan stok produk apa saja yang dapat disediakan untuk menunjang penjualan pada Agen stok HPAI.

4. Hasil dan Pembahasan

Data Mining secara umum terdiri atas 2 kata yaitu : data dan mining. Data merupakan kumpulan fakta yang bisa berupa objek atau kejadian yang tercatat dan terekam yang belum mempunyai arti (Minarni and Rahmad Hidayat 2013), sedangkan kata mining merupakan proses penambangan (Murlena and Syahindra 2020). Dengan demikian data mining dapat didefinisikan sebagai suatu proses penambangan data yang menghasilkan *output* berupa ilmu pengetahuan (Nofriansyah and Widi Nurcahyo 2015). Data mining merupakan bagian dari beberapa bidang ilmu pengetahuan seperti *machine learning*, pengenalan pola, statistik, basis data dan juga visualisasi dalam penanganan masalah untuk pengambilan informasi dari basis data yang besar (Larose 2015).

Algoritma C4.5 pada proses pemecahan masalah dan penemuan pengetahuan baru ada beberapa klasifikasi yang dilakukan secara umum diantaranya: Estimasi, Association, Klasifikasi, clustering dan Prediksi. Terdapat beberapa teknik klasifikasi yang dipergunakan sebagai solusi pemecahan masalah diantaranya : Algoritma C4.5, Algoritma *K-Nearest Neighbor*, ID3, *Naïve Bayesian Clasification*, CART (*Clasification And Regression Tree*) (Kusrini and Taufiq Luthfi 2009). Algoritma C4.5 adalah algoritma yang dipergunakan untuk membentuk pohon keputusan. Pada algoritma C4.5 untuk menyelesaikan masalah ada beberapa elemen yang diketahui yaitu : *Entropy* dan *Gain*. *Entropy* merupakan perkiraan jumlah bit yang dibutuhkan untuk bisa mengekstrak suatu kelas (negatif atau positif) dari sejumlah data acak pada ruang sampel S. *Gain* merupakan informasi yang diperoleh dari atribut A relative terhadap keluaran data S.

Stok merupakan barang dagang yang disimpan untuk kemudian dijual dalam operasi bisnis perusahaan. Dan bahan yang digunakan dalam proses produksi atau yang disimpan untuk tujuan itu. Stok adalah bahan atau barang yang disimpan yang akan digunakan untuk memenuhi tujuan tertentu.

Hasil dan Analisa Algoritma C4.5

Dalam hal ini penulis membahas mengenai prediksi ketersediaan stok barang. Untuk saat ini Produk HNI HPAI terdiri dari 3 jenis produk yaitu : *herbs Product, Health Food & Beverage dan Cosmetics & Home Care*. Dari tiga jenis produk inilah yang akan diprediksi produk mana yang tidak boleh kosong atau di stok lebih banyak dan produk apa yang harus dikurangi stoknya. Pada algoritma C4.5 ini langkah-langkah dalam menganalisa data-data yang harus dilakukan adalah :

1. Pemilihan atribut yang digunakan sebagai akar (node 1)
2. Membuat cabangnya untuk tiap-tiap nilai
3. Bagi kasus untuk tiap cabang
4. Lakukan langkah yang sama pada tiap cabang untuk semua kasus sehingga semua kelas nilainya sama.

Dalam menentukan atribut yang akan dipilih sebagai akar, ditentukan dari nilai *gain* yang tertinggi dari semua atribut-atribut yang ada. Tapi sebelum menghitung nilai dari *Gain* harus dihitung terlebih dulu nilai dari *Entropy*-nya. Rumus untuk menghitung nilai *Entropy* yaitu (Pritalia 2018) :

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

Keterangan :

S : himpunan Kasus

n : jumlah partisi S

p_i : proporsi dari S_i terhadap S

Untuk penghitungan nilai *Gain* dapat digunakan rumus berikut ini (Renhoran, Nurhandayani, and Septiana 2018) :

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

Keterangan :

S : himpunan kasus

A : Atribut

- n : jumlah partisi atribut A
 |Si| : jumlah kasus pada partisi ke-i
 |S| : jumlah kasus dalam S

Langkah pertama yang dilakukan yaitu menentukan Dataset seperti tabel 1 di bawah ini. Dalam menentukan dataset untuk produk HNI HPAI ini berdasarkan pada aplikasi online yang siap digunakan oleh setiap agenstok, dari aplikasi ini ditentukan jenis produknya (Herb Products, Cosmetics & Home Care dan Health Food & Beverage). Untuk harga, dikelompokkan berdasarkan kriteria : jika harga produk berkisar antara 10.000-35.000 dikategorikan Sangat Murah, jika harganya 35.000-100.000 dikategorikan Murah dan jika harga produk 100.000 keatas dikategorikan Mahal.

Sedangkan jumlah terjual dikelompokkan atas 2 bagian yaitu sedikit dan banyak. Untuk menentukan sedikit atau banyak berdasarkan atas : jika produk yang terjual dalam satu hari 1-2 termasuk Sedikit, jika produk terjual diatas 3 produk perhari maka termasuk Banyak. Keterangan Laris dan Kurang Laris ditentukan dari banyaknya jumlah terjual, jika jumlah terjual sedikit maka keterangannya kurang laris dan jika jumlah terjualnya banyak maka keterangannya pada dataset adalah laris.

Tabel 1. Data set Penjualan Produk HNI HPAI April 2022

No	Jenis Produk	Harga	Jml Terjual	Keterangan
1	<i>Herb Product</i>	Murah	Sedikit	Kurang Laris
2	<i>Herb Product</i>	Mahal	Banyak	Laris
3	<i>Herb Product</i>	Murah	Sedikit	Kurang Laris
4	<i>Cosmetics & Home Care</i>	Sangat Murah	Sedikit	Kurang Laris
5	<i>Cosmetics & Home Care</i>	Murah	Banyak	Laris
6	<i>Health Food & Beverage</i>	Murah	Banyak	Laris
7	<i>Cosmetics & Home Care</i>	Sangat Murah	Sedikit	Kurang Laris
8	<i>Herb Product</i>	Murah	Sedikit	Kurang Laris
9	<i>Herb Product</i>	Murah	Banyak	Laris
10	<i>Health Food & Beverage</i>	Mahal	Banyak	Kurang Laris
11	<i>Herb Product</i>	Murah	Sedikit	Kurang Laris
12	<i>Cosmetics & Home Care</i>	Sangat Murah	Sedikit	Kurang Laris
13	<i>Cosmetics & Home Care</i>	Mahal	Sedikit	Kurang Laris
14	<i>Health Food & Beverage</i>	Murah	Banyak	Laris
15	<i>Health Food & Beverage</i>	Murah	Banyak	Laris
16	<i>Health Food & Beverage</i>	Mahal	Sedikit	Kurang Laris
17	<i>Cosmetics & Home Care</i>	Murah	Sedikit	Kurang Laris
18	<i>Cosmetics & Home Care</i>	Murah	Sedikit	Kurang Laris
19	<i>Herb Product</i>	Murah	Sedikit	Kurang Laris
20	<i>Herb Product</i>	Murah	Sedikit	Kurang Laris

21	<i>Herb Product</i>	Mahal	Sedikit	Kurang Laris
22	<i>Herb Product</i>	Sangat Murah	Banyak	Laris
23	<i>Herb Product</i>	Sangat Murah	Sedikit	Kurang Laris
24	<i>Herb Product</i>	Sangat Murah	Banyak	Laris
25	<i>Cosmetics & Home Care</i>	Sangat Murah	Sedikit	Kurang Laris
26	<i>Cosmetics & Home Care</i>	Sangat Murah	Sedikit	Kurang Laris
27	<i>Cosmetics & Home Care</i>	Sangat Murah	Banyak	Laris
28	<i>Cosmetics & Home Care</i>	Sangat Murah	Sedikit	Kurang Laris
29	<i>Cosmetics & Home Care</i>	Sangat Murah	Banyak	Laris
30	<i>Cosmetics & Home Care</i>	Sangat Murah	Banyak	Laris
31	<i>Herb Product</i>	Murah	Sedikit	Kurang Laris
32	<i>Herb Product</i>	Mahal	Sedikit	Kurang Laris
33	<i>Herb Product</i>	Mahal	Sedikit	Kurang Laris
34	<i>Cosmetics & Home Care</i>	Sangat Murah	Banyak	Laris
35	<i>Cosmetics & Home Care</i>	Sangat Murah	Sedikit	Kurang Laris
36	<i>Cosmetics & Home Care</i>	Sangat Murah	banyak	Laris
37	<i>Health Food & Beverage</i>	Murah	Banyak	Laris
38	<i>Herb Product</i>	Murah	Banyak	Laris
39	<i>Health Food & Beverage</i>	Murah	Sedikit	Kurang Laris

Langkah kedua adalah menghitung total data terdiri atas (39) dan total data berdasarkan Keterangan “Kurang Laris” = 24 dan “Laris” = 15. Kemudian mengelompokkan dan menghitung semua total atribut disimpan dalam kolom “jumlah” dan total data atribut berdasarkan banyaknya penjualan, disimpan dalam kolom “Laris” dan “Kurang Laris”. Hasil pengelompokkan dapat dilihat pada tabel dibawah ini.

Tabel 2. Total Data, Atribut dan Kelas

Kategori	Atribut	Jumlah	Laris	Kurang Laris
Total		39	15	24
Jenis Produk				
	<i>Herbs Product</i>	16	5	11
	<i>Health food & beverage</i>	7	4	3
	<i>Cosmetic & home care</i>	16	6	10
Harga				
	Sangat Murah	15	7	8
	Murah	17	7	10
	Mahal	7	1	6
Jumlah Terjual				
	Sedikit	23	0	23
	Banyak	16	15	1

Langkah ke -3 yaitu mencari Entropy total dari data pada Tabel 2.

Nilai *Entropy Total*

$$\begin{aligned} \text{Entropy (Total)} &= \text{Entropy (S)} = \sum_{i=1} -p * \log_2 p_i \\ &= (-15/39 * \text{Log}_2 (15/39)) + (-24/39 * \text{Log}_2 (24/39)) = \\ &0,9612366 \end{aligned}$$

Langkah ke 4 yaitu : mencari *Entropy* kategori (Jenis produk, harga dan jumlah terjual) dari data Tabel 2. Masing-masing atribut dihitung berdasarkan berdasarkan jumlah dan kelas (keterangan Laris atau Kurang Laris).

$$\begin{aligned} \text{Entropy (Jenis Produk, Herbs Product)} &= (-5/16 * \text{Log}_2 (5/16)) + (-11/16 * \text{Log}_2 \\ &(11/16)) = 0,89603823 \end{aligned}$$

$$\begin{aligned} \text{Entropy (Jenis Produk, Health Food \& Beverage)} &= (-3/7 * \text{Log}_2 (3/7)) + (-4/7 * \text{Log}_2 \\ &(4/7)) = 0,98522814 \end{aligned}$$

$$\begin{aligned} \text{Entropy (Jenis Produk, Cosmetics \& Home Care)} &= (-6/16 * \text{Log}_2(6/16)) + (-10/16 * \\ &\text{Log}_2(10/16)) = 0,954434 \end{aligned}$$

$$\begin{aligned} \text{Entropy (Harga, Sangat murah)} &= (-7/15 * \text{Log}_2 (7/15)) + (-8/15 * \text{Log}_2 (8/15)) = \\ &0,99679163 \end{aligned}$$

$$\begin{aligned} \text{Entropy (Harga, Murah)} &= (-7/17 * \text{Log}_2 (7/17)) + (-10/17 * \text{Log}_2 (10/17)) = \\ &0,97741782 \end{aligned}$$

$$\begin{aligned} \text{Entropy (Harga, Mahal)} &= (1/7 * \text{Log}_2 (1/7)) + (-6/7 * \text{Log}_2 (6/7)) = \\ &0,59167278 \end{aligned}$$

$$\begin{aligned} \text{Entropy (Jumlah Terjual, Sedikit)} &= (-0/23 * \text{Log}_2 (0/23)) + (-23/23 * \text{Log}_2 \\ &(23/23)) = 0 \end{aligned}$$

$$\begin{aligned} \text{Entropy (Jumlah Terjual, Banyak)} &= (-15/16 * \text{Log}_2 (15/16)) + (-1/16 * \text{Log}_2 \\ &(1/16)) = 0,33729007 \end{aligned}$$

Langkah ke 5 yaitu menentukan nilai *gain* masing-masing kategori (jenis produk, harga dan jumlah terjual) dan memilih nilai *gain* tertinggi.

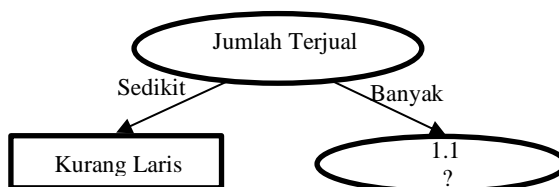
$$\begin{aligned} \text{Gain (Jenis Produk)} &= (0,9612366) - (16/39 * 0,89603823) - (7/39 * 0,98522814) - \\ &(16/39 * 0,954434) \\ &= 0,0252327 \end{aligned}$$

$$\begin{aligned} \text{Gain (Harga)} &= (0,9612366) - (15/39 * 0,99679163) - (17/39 * 0,97741782) - \\ &(7/39 * 0,59167278) \\ &= 0,0456036 \end{aligned}$$

$$\begin{aligned} \text{Gain (jumlah terjual)} &= (0,9612366) - (23/39 * 0) - (16/39 * 0,33729007) \\ &= 0,8228612 \end{aligned}$$

Tabel 3. Perhitungan Node 1

Node	Kategori	Atribut	Jml	Laris	Kurang Laris	Entropy	Gain
1	Total		39	15	24	0,9612366	
	Jenis Produk						0,0252327
		<i>Herbs Product</i>	16	5	11	0,89603823	
		<i>Health food & beverage</i>	7	4	3	0,98522814	
		<i>Cosmetic & home care</i>	16	6	10	0,954434	
	Harga						0,0456036
		Sangat Murah	15	7	8	0,99679163	
		Murah	17	7	10	0,97741782	
		Mahal	7	1	6	0,59167278	
	Jumlah Terjual						0,8228612
		Sedikit	23	0	23	0	
		Banyak	16	15	1	0,33729007	



Gambar 2. Pohon Keputusan Hasil Perhitungan Node 1

Tabel 4. Perhitungan Node 1.1

Node	Kategori	Atribut	Jumlah	Laris	Kurang Laris	Entropy	Gain
1.1	Jumlah terjual-banyak		16	15	1	0,33729007	
	Jenis Produk						0,1116875
		<i>Herbs Product</i>	5	5	0	0	
		<i>Health food & beverage</i>	5	4	1	0,72192809	
		<i>Cosmetic & home care</i>	6	6	0	0	

	Harga						0,212290 1
		Sangat Murah	7	7	0	0	
		Murah	7	7	0	0	
		Mahal	2	1	1	1	

ENTROPY

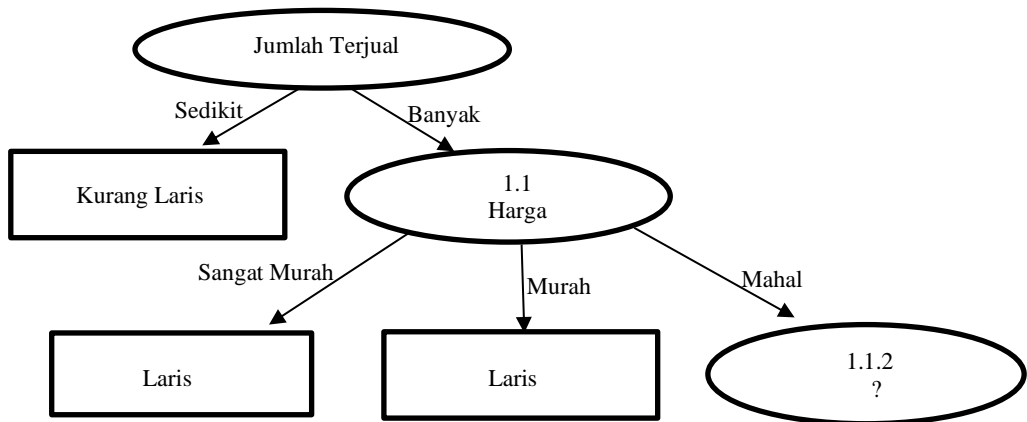
Entropy (jml terjual-banyak, Herb product) :
 $= (-4/5 * \text{Log}_2(4/5)) + (-1/5 * \text{Log}_2(1/5)) = 0,72192809$

Entropy (jml terjual-banyak, Mahal) :
 $= ((-1/2 * \log_2(1/2)) + ((-1/2 * \text{Log}_2(1/2))) = 1$

GAIN

Gain (jml terjual-byk, jenis produk)
 $= ((0,33729007) - ((5/16 * 0) + (5/16 * 0,72192809) + (6/16 * 0))) = 0,1116875$

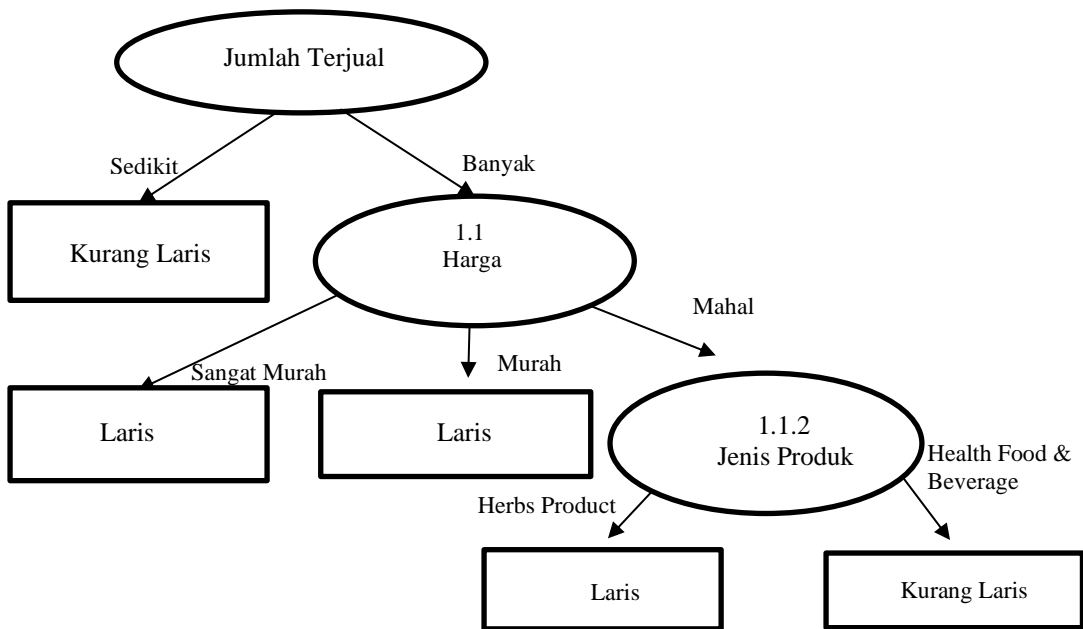
Gain (jml terjual-byk, harga)
 $= ((0,33729007) - ((7/16 * 0) + (7/16 * 0) + (2/16 * 1))) = 0,2122901$



Gambar 3. Pohon Keputusan Hasil Perhitungan Node 1.1

Tabel 5. Perhitungan Node 1.1.2

Node	Kategori	Atribut	Jml	Laris	Kurang Laris	Entrop y	Gain
1.1.2	Jumlah terjual-Banyak dan Harga-Mahal		2	1	1	1	
	Jenis Produk						1
		Herbs Product	1	1	0	0	
		Health food & beverage	1	0	1	0	
		Cosmetics & Home Care	0	0	0	0	



Gambar 4. Pohon Keputusan Hasil Perhitungan *Node 1.1.2*

5. Kesimpulan

Berdasarkan uraian yang telah dilakukan pada analisa diatas, dapat diambil kesimpulan bahwa berdasarkan analisa yang dilakukan terhadap produk HNI HPAI yang jumlah terjualnya banyak meskipun harganya tergolong mahal adalah produk herbal (*herbs product*). Sesuai dengan pohon keputusan hasil perhitungan node 1.1.2 maka produk herbal (*herbs product*) yang harus di stok lebih banyak dibandingkan dengan health food and beverage. Sehingga dengan penerapan data mining menggunakan algoritma C4.5 ini maka para agenstok dapat terhindar dari penumpukan stok yang kurang laris.

Data-data penjualan dan pembelian pada Business Center (BC) tentu lebih banyak lagi dibandingkan data pada Agenstok, maka penerapan algoritma C4.5 ini juga bisa dikembangkan pada Business Center (BC) dalam menentukan stok barang khususnya pada produk herbalnya yang laris berdasarkan mereknya (nama-nama produk herbal) agar sesuai dengan kebutuhan Agenstok atau konsumen.

6. Ucapan Terima Kasih

Dalam penyusunan penelitian ini, peneliti mengucapkan terima kasih kepada pihak-pihak yang terlibat dalam pengumpulan data, terutama kepada ibu desmita di Padang, Sumatera Barat yang merupakan merupakan AgenStok produk HPAI.

7. Pernyataan Penulis

Penulis menyatakan bahwa tidak ada konflik kepentingan terkait publikasi artikel ini. Penulis menegaskan bahwa data dan makalah bebas dari plagiarisme.

Bibliografi

- Esthiningtyas, G., & Prasetyaningrum, P. T. (2020). Penerapan Algoritma C4. 5 Untuk Menentukan Persediaan Obat (Studi Kasus Di RS Bethesda Yogyakarta). *Journal Of Information System And Artificial Intelligence*, 1(1), 25-33.
- Fikri, Abdul, and Wiwi Verina. 2020. "Penerapan Data Mining Untuk Prediksi Penjualan Alat Medis Menggunakan Algoritma C4.5 PT. Murni Indah Sentosa." *Infosys (Information System) Journal* 5 (1): 70–83. <https://doi.org/10.22303/infosys.5.1.2020.70-83>.
- Hayati, N. (2020). Optimalisasi Prediksi Penjualan Produk Herbal Menggunakan Metode Monte Carlo dalam Meningkatkan Transaksi. *Jurnal Informatika Ekonomi Bisnis*, 117-122.
- Jumiyati, and Putri Taqwa Prasetyaningrum. 2020. "Penerapan Algoritma C4.5 Untuk Mengklasifikasi Hasil Produksi Kunir Putih (Studi Kasus CV Windra Mekar)." *Journal Of Information System And Artificial Intelligence* 1 (1): 9–16. <http://jisai.mercubuana-yogya.ac.id/index.php/jisai/article/view/30>.
- Kusrini, Kusrini, and Emha Taufiq Luthfi. 2009. *Algoritma Data Mining*. Yogyakarta: Andi Offset.
- Larose, Daniel T. 2015. *Data Mining and Predictive Analytics*. John Wiley & Sons.
- Minarni, S. Si, and S. T. Rahmad Hidayat. 2013. "Rancang Bangun Aplikasi Sistem Pakar Untuk Kerusakan Komputer Dengan Metode Backward Chaining." *Jurnal Teknoif Teknik Informatika Institut Teknologi Padang* 1 (1): 26–35. <https://doi.org/10.21063/jtif.2013.V1.1.26-35>.
- Murlena, Murlena, and Wandu Syahindra. 2020. "Data Mining Pengolahan Penempatan Library Books Menggunakan Metode Association Rule Dengan Algoritma Apriori." *Jurnal INSTEK (Informatika Sains Dan Teknologi)* 5 (2): 199–208. <https://doi.org/10.24252/instek.v5i2.16203>.
- Nofriansyah, Dicky, and Gunadi Widi Nurcahyo. 2015. *Algoritma Data Mining Dan Pengujian*. Vol. viii. Yogyakarta: Deepublish.
- Pritalia, Generosa Lukhayu. 2018. "Penerapan Algoritma C4.5 Untuk Penentuan Ketersediaan Barang E-Commerce." *Indonesian Journal of Information Systems* 1 (1): 47–56. <https://doi.org/10.24002/ijis.v1i1.1727>.
- Renhoran, Bahrin Said, Nova Nurhandayani, and Laila Septiana. 2018. "Penerapan Algoritma C4.5 Untuk Menentukan Data Stok Dan Target Permintaan Material

Yang Paling Dibutuhkan Gudang Logistik Pada Pt.Pln (Persero) Area Kebon Jeruk.” *INTI Nusa Mandiri* 12 (2): 13–20.
<https://doi.org/10.33480/inti.v12i2.1564>.

Widyastuti, Rini, Karmila Suryani, Ade Fitri Rahmadani, Triadmoko Denny Fatrosa, and Wandi Syahindra. 2021. “Matrix Integration with Jitsi Conference Server for Online Learning.” *Knowbase: International Journal of Knowledge in Database* 1 (2): 106–15.
<https://doi.org/10.30983/ijokid.v1i2.5040>.