

Penerapan Data Mining Untuk Klasifikasi Data Penjualan Pakan Ternak Terlaris Dengan Algoritma C4.5

Dandi Muhamad Musa¹⁾, Dimas Sakti²⁾, Keiko Angel Shantiony³⁾, Sandi Kurnia Putri Zega⁴⁾, Sevtyan Hamzah⁵⁾, Yohanes Julfani Zega⁶⁾, Baginda Oloan Lubis^{7)*}

¹⁾²⁾³⁾⁴⁾⁵⁾⁶⁾⁷⁾Program Studi Sistem Informasi, Fakultas Teknik dan Informatika,
Universitas Bina Sarana Informatika

^{*)}Correspondence Author: baginda.bio@bsi.ac.id, Jakarta, Indonesia

DOI: <https://doi.org/10.37012/jtik.v10i1.1985>

Abstrak

Penelitian ini fokus pada tantangan terkait data penjualan di PT Fresh X. Masalah utamanya adalah kesulitan dalam mendapatkan informasi strategis seperti tingkat penjualan per periode dan produk pakan ternak yang paling diminati. Untuk mengatasi masalah ini, dilakukan analisis *data mining* pada *dataset* penjualan PT Fresh X. Ketersediaan data yang melimpah memberikan peluang untuk memperoleh informasi yang diperlukan, dengan harapan dapat menjadi dasar untuk pengambilan keputusan yang efektif dalam merancang solusi bisnis. Penerapan teknik *data mining* diharapkan dapat mempercepat proses pengambilan keputusan dan memungkinkan perusahaan mengubah informasi dari data transaksi menjadi pengetahuan baru, tanpa harus langsung menggunakan *data warehouse* sebagai sumber *data mining*. Tujuan penelitian ini adalah menggunakan algoritma C4.5 untuk mengklasifikasikan penjualan pakan ternak menjadi kategori yang laris atau tidak laris, serta mengukur tingkat akurasi algoritma C4.5 dalam melakukan klasifikasi tersebut. Tahapan penelitian meliputi, Pengolahan data, Transformasi data, Penerapan algoritma C4.5 (meliputi pencarian nilai *entropy* dan *gain*), dan Pengolahan data dengan *decision tree*. *Decision tree* mengelompokkan berbagai jenis pakan ternak berdasarkan kategori, harga, kuantitas, dan informasi lainnya. Diagram ini bermanfaat untuk memahami karakteristik dan perbandingan antar jenis pakan. Hasil penelitian menunjukkan *gain* tertinggi terdapat pada kategori pakan dengan nilai 0.306739968 dan *entropy* pakan ayam pedaging dengan nilai 0.99107606. Hal ini menunjukkan bahwa pakan ayam pedaging merupakan produk paling laris berdasarkan hasil pengolahan data.

Kata Kunci: Pakan Ternak, Klasifikasi, Algoritma C4.5

Abstract

This research focuses on challenges related to sales data at PT Fresh X. The main problem is the difficulty in obtaining strategic information such as sales levels per period and the most popular animal feed products. To overcome this problem, data mining analysis was carried out on the sales dataset at PT Fresh X. The availability of abundant data provides opportunities to obtain the necessary information, with the hope that it can become the basis for making effective decisions in designing business solutions. The application of data mining techniques is expected to speed up the decision-making process and enable companies to convert information from transaction data into new knowledge, without having to directly use a data warehouse as a data mining source. The aim of this research is to use the C4.5 algorithm to classify animal feed sales into categories that are selling well or not selling well, as well as measuring the level of accuracy of the C4.5 algorithm in carrying out this classification. Research stages include, data processing, data transformation, application of the C4.5 algorithm (including searching for entropy and gain values), and data processing with decision trees. Decision trees group various types of animal feed based on category, price, quantity and other information. This diagram is useful for understanding the characteristics and comparisons between types of feed. The research results showed that the highest gain was in the feed category with a value of 0.306739968 and the entropy of broiler feed with a value of 0.99107606. This shows that broiler chicken feed is the best-selling product based on the results of data processing.

Keywords: Animal Feed, Classification, C4.5 Algorithm

PENDAHULUAN

Penjualan pakan ternak memiliki peran krusial dalam industri peternakan dan memberikan dampak signifikan terhadap pertumbuhan ekonomi serta ketahanan pangan suatu negara. Pakan ternak menjadi elemen utama dalam pemeliharaan hewan ternak, dan tingkat penjualan dapat dipengaruhi oleh kepuasan pembeli. Kualitas dan ketersediaan pakan secara langsung berdampak pada produktivitas dan kesejahteraan ternak. Sayangnya, data yang seharusnya menjadi sumber wawasan lebih mendalam seringkali hanya digunakan untuk menyusun laporan penjualan bulanan tanpa mengeksplorasi informasi atau pengetahuan baru dari pola transaksi yang ada. (Lalo et al., 2021).

Kemajuan pesat dalam teknologi informasi telah tersebar luas di berbagai sektor kehidupan, termasuk industri peternakan modern. Saat ini, industri peternakan menghadapi sejumlah tantangan, terutama dalam konteks penjualan pakan ternak. Dengan pertumbuhan terus-menerus dalam jumlah penduduk manusia, permintaan terhadap produk hewani seperti daging, susu, dan telur juga terus meningkat. Inilah yang mendorong para peternak untuk meningkatkan produksi ternak mereka dengan memastikan pasokan pakan yang memadai dan berkualitas. (Lia & Setyo Wibagso, 2022).

Disaat saat yang sama, dinamika industri pakan ternak dipengaruhi oleh perubahan iklim, fluktuasi harga komoditas, dan perubahan pola konsumsi masyarakat. (Pritalia, 2018). Tantangan ini dapat diatasi dengan melakukan penelitian dan inovasi yang meliputi formulasi pakan, manajemen rantai pasokan, dan teknologi produksi.

Dengan mempertimbangkan konteks ini, penelitian ini melakukan penyelidikan dan analisis terhadap tren, faktor-faktor, dan dampak yang berkaitan dengan penjualan pakan ternak (Diansyah & Exprada, 2022). Akan dibahas berbagai aspek dalam hal ini, antara lain:

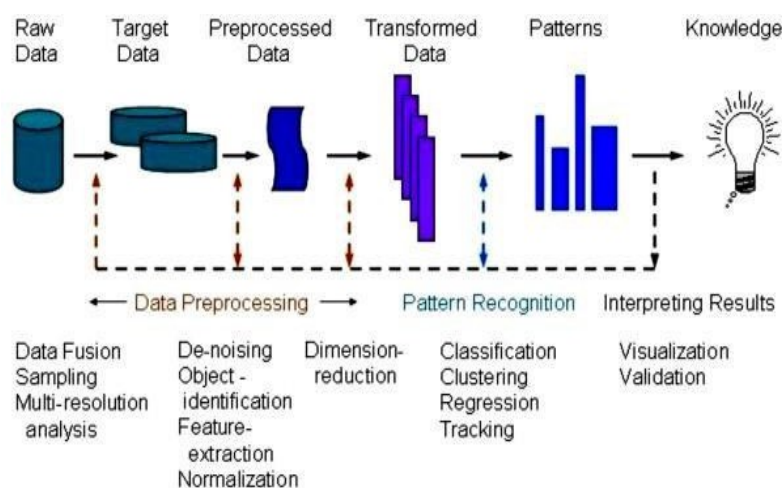
1. Faktor-Faktor yang Mempengaruhi Penjualan Pakan Ternak: Apa saja yang menentukan permintaan dan penawaran pakan ternak? Bagaimana kondisi ekonomi, perubahan iklim, dan inovasi teknologi mempengaruhi penjualan pakan ternak?
2. Perkembangan dalam Formulasi Pakan: Bagaimana formulasi pakan yang lebih baik dapat meningkatkan produktivitas ternak? Apa saja inovasi terkini dalam pakan ternak yang dapat memperbaiki efisiensi produksi?

3. Pengaruh Penjualan Pakan Ternak: Bagaimana penjualan pakan ternak berdampak pada industri peternakan secara umum? Bagaimana peningkatan penjualan pakan ternak dapat mendukung ketahanan pangan dan pembangunan ekonomi?
4. Kesejahteraan Hewan dalam Penjualan Pakan Ternak: Bagaimana kesejahteraan hewan berkaitan dengan penjualan pakan ternak? Bagaimana pakan yang berkualitas dapat memperbaiki kesehatan dan kesejahteraan ternak?

Penelitian ini bertujuan untuk mengkaji lebih dalam dinamika industri penjualan pakan ternak di PT Fresh X, serta peran pentingnya dalam mendukung keberlanjutan dan perkembangan industri peternakan. Hasil penelitian ini diharapkan akan memberikan pengetahuan yang berharga bagi para pemangku kepentingan di industri peternakan dan memudahkan pengambilan keputusan yang lebih baik dalam manajemen dan investasi di sektor ini. (Abdullah et al., 2022)

METODE

Penambangan data adalah ekstraksi informasi yang berarti dari sejumlah besar data. Ini digunakan di banyak aplikasi seperti data pendidikan, web, dan penambangan teks. Semua aktivitas ini dilakukan secara otomatis dan dapat ditemukan dengan cepat oleh pemrogram. Jenis utama penambangan data meliputi klasifikasi, pengelompokan, dan aturan asosiasi.



Sumber: (Anggraini et al., 2018)

Gambar 1. Tahapan Proses Data Mining

Proses penambangan data mencakup proses yang disebut klasifikasi. Klasifikasi melibatkan penemuan model yang mendeskripsikan kelas data. Pada proses klasifikasi, data dimuat dalam bentuk *decision tree* (pohon keputusan).

Iterative Dychotomizer Version 3 (ID3) adalah salah satu pendekatan yang sangat terkenal dalam metode pohon keputusan. Algoritma ini adalah pembelajaran pohon keputusan yang melakukan pencarian secara rakus (*greedy*), yang berarti hasilnya mungkin tidak selalu mencapai optimal. Algoritma ini dikembangkan oleh Ross Quinlan pada tahun 1986 dan termasuk dalam kategori algoritma pembelajaran mesin berbasis aturan. ID3 berusaha untuk membangun model klasifikasi berbentuk pohon keputusan dengan pendekatan *top-down* (dari atas ke bawah). Proses ini melibatkan evaluasi semua atribut dengan menggunakan ukuran statistik, biasanya *information gain*, untuk mengukur sejauh mana setiap atribut efektif dalam mengklasifikasikan sekumpulan sampel data. (Iriadi et al., 2020).

Klasifikasi adalah menemukan model dari kumpulan data pelatihan yang membedakan atribut ke dalam kategori atau kelas yang sesuai. (Azwanti, 2018). Klasifikasi memegang peranan yang sangat penting dalam proses algoritma C4.5 menghasilkan model pohon atau aturan yang mudah diinterpretasikan dan dikonversi ke struktur aturan SQL (bahasa kueri), memiliki tingkat akurasi yang dapat diterima, dapat menangani atribut diskrit dan numerik, dan sangat efisien dalam pemrosesan. (Suyanto, 2017). Tujuan utama klasifikasi adalah mengidentifikasi pola dalam data dan memprediksi kelas objek yang tidak diketahui berdasarkan informasi dari objek yang diketahui. Dalam penambangan data dan pembelajaran mesin, ini adalah tugas umum. Ini melibatkan pengembangan model atau algoritma yang dapat mempelajari pola dari data pelatihan yang sudah diberi label dan mengklasifikasikan data baru ke dalam kategori yang sesuai.

Algoritma C4.5 memiliki banyak aplikasi di berbagai bidang. Misalnya, dalam dunia komputer, algoritma digunakan untuk mengembangkan perangkat lunak, mesin pencarian web, dan kecerdasan buatan. (Dwiasnati et al., 2023). Algoritma C4.5 adalah algoritma pembelajaran mesin yang digunakan untuk membuat pohon keputusan dari data. Algoritma ini merupakan pengembangan lebih lanjut dari algoritma ID3. (Suntoro, 2019). Cara kerja

algoritma C4.5 adalah dengan membentuk pohon keputusan yang menghasilkan keputusan. (Iriadi & Nuraeni, 2016). Algoritma ini terdiri dari beberapa tahap, yaitu:

- a. **Pemilihan atribut *root*:** Tahap ini bertujuan untuk memilih atribut yang paling informatif untuk membagi data menjadi kelompok-kelompok yang lebih homogen. Atribut *root* adalah atribut yang ditempatkan di simpul paling atas dari pohon keputusan.
- b. **Pemisahan *dataset*:** Tahap ini bertujuan untuk membagi data menjadi subset-subset berdasarkan nilai-nilai dari atribut *root*. Setiap subset akan menjadi cabang dari simpul *root*, dan akan menjadi input untuk tahap selanjutnya.
- c. **Penghitungan entropi:** Tahap ini bertujuan untuk mengukur seberapa beragam data dalam setiap subset. Entropi adalah ukuran ketidakpastian atau ketidakteraturan dari data. Entropi semakin tinggi jika data semakin heterogen, dan semakin rendah jika data semakin homogen. Entropi dapat dihitung dengan rumus berikut:

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

di mana $E(S)$ adalah entropi dari subset S , n adalah jumlah kelas yang ada dalam data, dan p_i adalah proporsi data yang termasuk dalam kelas i .

- d. **Penghitungan *gain*:** Tahap ini bertujuan untuk mengukur seberapa besar pengurangan entropi yang dapat dicapai dengan memilih suatu atribut untuk membagi data. *Gain* adalah selisih antara entropi sebelum dan sesudah pemisahan data. *Gain* dapat dihitung dengan rumus berikut:

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

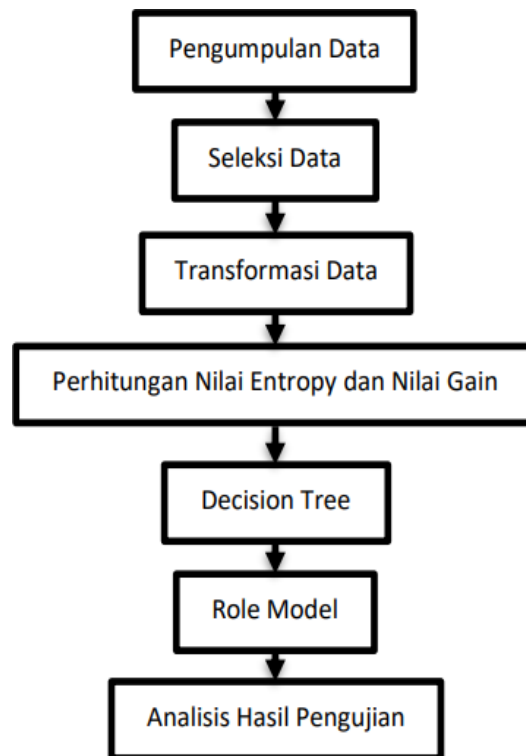
di mana $G(S,A)$ adalah *gain* dari subset S dengan menggunakan atribut A , $|S_i|$ adalah subset dari S yang memiliki nilai i untuk atribut A , dan $|S|$ adalah jumlah data dalam subset S .

- e. **Pemilihan atribut selanjutnya:** Tahap ini bertujuan untuk memilih atribut yang memiliki *gain* tertinggi untuk menjadi simpul selanjutnya dalam pohon keputusan. Atribut yang dipilih akan menjadi dasar untuk membagi subset-subset yang telah dibentuk sebelumnya menjadi subset-subset yang lebih kecil lagi. Proses ini diulangi sampai semua data memiliki kelas yang sama atau tidak ada atribut yang tersisa.

f. Pembentukan pohon keputusan: Tahap ini bertujuan untuk membentuk pohon keputusan yang dapat merepresentasikan hubungan antara atribut-atribut dan kelas-kelas dalam data. Pohon keputusan terdiri dari simpul-simpul dan cabang-cabang. Simpul-simpul merepresentasikan atribut-atribut yang digunakan untuk membagi data, dan cabang-cabang merepresentasikan nilai-nilai dari atribut-atribut tersebut. Simpul-simpul yang tidak memiliki cabang disebut sebagai simpul daun, dan merepresentasikan kelas-kelas yang ada dalam data. Pohon keputusan dapat digunakan untuk melakukan klasifikasi atau prediksi terhadap data baru dengan mengikuti cabang-cabang yang sesuai dengan nilai-nilai atribut data tersebut.

Analisis data merupakan proses evaluasi dan pengolahan data untuk memperoleh informasi penting, merumuskan kesimpulan, dan membantu dalam proses pengambilan keputusan. Memiliki keterampilan dalam analisis data adalah suatu keharusan bagi seorang praktisi data. Proses ini memerlukan pemikiran kritis dan kemampuan pemecahan masalah yang baik. Keterampilan ini penting untuk menentukan metode analisis data yang tepat. (Kodari et al., 2020). Metode analisis yang benar dapat memberikan dampak signifikan pada hasilnya. Analisis data dilakukan dengan mengumpulkan data penjualan pakan ternak selama 3 bulan terakhir. Selain itu, data penjualan pakan ternak juga digunakan sebagai data analisis untuk memprediksi jumlah penjualan pakan ternak di masa mendatang, sesuai dengan klasifikasi dan merek pakan ternak yang ada. Proses analisis yang berjalan dengan baik akan menghasilkan informasi yang jelas dan cepat. (Salim et al., 2022)

Tahapan penelitian adalah seperangkat langkah, prosedur, dan teknik perencanaan, pelaksanaan, dan analisis penelitian yang bertujuan untuk memperoleh informasi atau untuk memahami fenomena secara sistematis. Metode penelitian membantu peneliti menggambarkan langkah-langkah yang diambilnya untuk menjawab pertanyaan penelitian atau mencapai tujuan penelitian. (Cynthia & Ismanto, 2018). Berikut ini adalah tahapan proses penelitian:



Sumber: Hasil Penelitian (2024)

Gambar 2. Tahapan Penelitian

Tahapan penelitian dapat dijelaskan setiap tahapnya sebagai berikut:

1. Pengumpulan Data

Pengumpulan data merupakan tahap penting dalam penelitian yang bertujuan untuk mendapatkan informasi yang akan digunakan. Tidak ada aturan baku mengenai jumlah sampel yang harus diambil untuk mewakili populasi. Namun, secara umum, semakin besar sampelnya, semakin besar kemungkinan representatif terhadap populasi. (Lubis, 2014). Data yang dikumpulkan berfokus pada penjualan pakan ternak selama periode tiga bulan.

2. Seleksi Data

Seleksi data merujuk pada data yang diterapkan dalam proses algoritma klasifikasi C4.5. Tujuan dari pemilihan data ini adalah menciptakan kumpulan data target, memilih subset data, atau menfokuskan pada subkumpulan variabel atau sampel data yang akan diidentifikasi. (Arifin & Fitriana, 2018)

3. Transformasi Data

Transformasi data dilakukan terutama untuk mengubah skala pengukuran data asli ke format lain agar data memenuhi asumsi yang mendasari analisis varians dan konsisten dengan perhitungan algoritma C4.5.

4. Perhitungan *Entropy* dan *Information Gain*

Hitung *entropi* semua atribut/variabel, *entropi* pada persamaan (1), dan perolehan informasi pada persamaan (2) untuk mencari perolehan informasi tertinggi, yang digunakan sebagai simpul akar saat membangun pohon keputusan.

5. Pohon Keputusan (*Decision Tree*)

Pohon keputusan merupakan hasil proses perhitungan *entropy* dan *information gain* setelah setiap atribut pohon mempunyai kelas dan proses perhitungan tersebut diulangi hingga tidak dapat dilakukan lagi.

6. Aturan-aturan/*Rule Model*

Model aturan/*rule* adalah deskripsi deskriptif dari pohon keputusan.

7. Analisis Hasil Pengujian

Analisis hasil pengujian adalah proses menafsirkan dan mengevaluasi data hasil percobaan dan tes untuk memperoleh pemahaman lebih dalam terhadap suatu fenomena atau masalah penelitian. Analisis pada penelitian ini merupakan suatu analisis data yang dipergunakan apabila kesimpulan-kesimpulan yang diperoleh berupa angka-angka dan juga dalam perhitungan dipergunakan rumus yang ada hubungannya dengan analisis penulisan. (Salim & Lubis, 2017)

HASIL DAN PEMBAHASAN

1. Pengolahan Data

Berdasarkan tahapan penelitian dilakukan pengumpulan data untuk diolah. Pengolahan data merupakan manipulasi data sehingga dapat digunakan untuk analisis, pengambilan keputusan, atau penyajian informasi yang berguna. Berikut ini data yang digunakan dalam penelitian ini:

Tabel 1. Data Penjualan Pakan Ternak Januari Sampai Maret 2023

No	Nama Barang	Kategori	Harga/kg	Quantity			Total Harga		
				Jan	Feb	Mar	Januari	Februari	Maret
1	511 Bravo	Pakan Ayam Pedaging	Rp11,000	21	8	8	231,000	88,000	88,000
2	594 Bangkok	Pakan Ayam Pedaging	Rp10,500		50		-	525,000	-
3	B12L New hope	Pakan Ayam Pedaging	Rp6,500	1642	758	450	10,673,000	4,927,000	2,925,000
4	B12L New hope Pelet	Pakan Ayam Pedaging	Rp6,500		150	100	-	975,000	650,000
5	Beras Merah	Pakan Unggas	Rp11,000			25	-	-	275,000
6	BR11	Pakan Ayam Pedaging	Rp10,500			200	-	-	2,100,000
7	BR12	Pakan Ayam Pedaging	Rp10,500		50	150	-	525,000	1,575,000
8	Cargil Petelur	Pakan Ayam Petelur	Rp8,500		50	300	-	425,000	2,550,000
9	Dedak Bekatul	Pakan Unggas	Rp5,000	16	9	33	80,000	45,000	165,000
10	Dedak Campur	Pakan Unggas	Rp5,000		7		-	35,000	-
11	Egg Max Petelur	Pakan Ayam Petelur	Rp8,500	211	206	258	1,793,500	1,751,000	2,193,000
12	Millet Kuning	Pakan Unggas	Rp15,000		9	3	-	135,000	45,000
13	New Hope D90	Pakan Bebek	Rp8,500		50	50	-	425,000	425,000
14	Omega -3	Pakan Ikan	Rp8,500	99	76	80	841,500	646,000	680,000
15	P02 New Hope	Pakan Unggas	Rp9,500	11	38	70	104,500	361,000	665,000
16	Wonokoyo Buras	Pakan Ayam Pedaging	Rp6,500	428	916	554	2,782,000	5,954,000	3,601,000
17	Wonokoyo M-3	Pakan Ikan	Rp8,500	11			93,500	-	-
18	Wonokoyo BR1	Pakan Ayam Pedaging	Rp9,000	2050	1200	3398	18,450,000	10,800,000	30,582,000
19	Wonokoyo BR1 88	Pakan Ayam Pedaging	Rp8,500			3500	-	-	29,750,000
20	Wonokoyo Petelur	Pakan Ayam Petelur	Rp8,500	193	4	2	1,640,500	34,000	17,000

Sumber: Hasil Penelitian (2024)

2. Transformasi Data

Berikut adalah hasil dari transformasi data penjualan yang akan diolah:

Tabel 2. Hasil Transformasi Data Penjualan

No	Nama Barang	Kategori	Harga	Quantity	Keterangan
1	511 Bravo	Pakan Ayam Pedaging	Mahal	Sedikit	Tidak Laris
2	594 Bangkok	Pakan Ayam Pedaging	Mahal	Sedikit	Tidak Laris
3	B12L New hope	Pakan Ayam Pedaging	Murah	Banyak	Laris
4	B12L New hope Pelet	Pakan Ayam Pedaging	Murah	Sedikit	Tidak Laris
5	Beras Merah	Pakan Unggas	Mahal	Sedikit	Tidak Laris
6	BR11	Pakan Ayam Pedaging	Mahal	Sedikit	Tidak Laris
7	BR12	Pakan Ayam Pedaging	Mahal	Sedikit	Tidak Laris
8	Cargil Petelur	Pakan Ayam Petelur	Murah	Sedikit	Tidak Laris
9	Dedak Bekatul	Pakan Unggas	Murah	Sedikit	Tidak Laris
10	Dedak Campur	Pakan Unggas	Murah	Sedikit	Tidak Laris
11	Egg Max Petelur	Pakan Ayam Petelur	Murah	Banyak	Laris
12	Millet Kuning	Pakan Unggas	Mahal	Sedikit	Tidak Laris
13	New Hope D90	Pakan Bebek	Murah	Banyak	Laris
14	Omega -3	Pakan Ikan	Murah	Banyak	Laris
15	P02 New Hope	Pakan Unggas	Murah	Banyak	Laris
16	Wonokoyo Buras	Pakan Ayam Pedaging	Murah	Banyak	Laris
17	Wonokoyo M-3	Pakan Ikan	Murah	Sedikit	Tidak Laris
18	Wonokoyo Pedaging BR1	Pakan Ayam Pedaging	Murah	Banyak	Laris
19	Wonokoyo Pedaging BR1 88	Pakan Ayam Pedaging	Murah	Banyak	Laris
20	Wonokoyo Petelur	Pakan Ayam Petelur	Murah	Sedikit	Tidak Laris

Sumber: Hasil Penelitian (2024)

3. Penerapan Algoritma C4.5

Implementasi Algoritma C4.5 mencakup serangkaian langkah seperti pemilihan *atribut root*, pemisahan *dataset*, penghitungan *entropi*, dan penghitungan *gain*. *Atribut root* dipilih sebagai dasar untuk membagi *dataset*, dan proses ini diulangi sampai terbentuk pohon keputusan. Data dibagi berdasarkan nilai atribut, dan *entropi* dihitung untuk menilai keragaman data. *Gain* dihitung untuk setiap atribut, dan atribut dengan *gain* tertinggi dipilih. Proses ini berlanjut sampai pohon keputusan terbentuk yang dapat digunakan untuk klasifikasi atau prediksi.

a) Mencari Nilai *Entropy*

1. *Entropy* Total

$$\begin{aligned} \text{Entropy [Total]} &: \left(-\frac{8}{20} \times \log_2 \left(\frac{8}{20}\right)\right) + \left(-\frac{12}{20} \times \log_2 \left(\frac{12}{20}\right)\right) \\ &: 0,970950594 \end{aligned}$$

2. *Entropy* kategori pakan

$$\begin{aligned} - \text{Entropy [Pakan Ayam Pedaging]} &: \left(-\frac{4}{9} \times \log_2 \left(\frac{4}{9}\right)\right) + \\ &\left(-\frac{5}{9} \times \log_2 \left(\frac{5}{9}\right)\right) \\ &: 0,991076059 \\ - \text{Entropy [Pakan Ayam Petelur]} &: \left(-\frac{1}{3} \times \log_2 \left(\frac{1}{3}\right)\right) + \left(-\frac{2}{3} \times \log_2 \left(\frac{2}{3}\right)\right) \\ &: 0,918295834 \\ - \text{Entropy [Pakan Unggas]} &: \left(-\frac{1}{5} \times \log_2 \left(\frac{1}{5}\right)\right) + \left(-\frac{4}{5} \times \log_2 \left(\frac{4}{5}\right)\right) \\ &: 0,721928094 \\ - \text{Entropy [Pakan Bebek]} &: \left(-\frac{1}{1} \times \log_2 \left(\frac{1}{1}\right)\right) + \left(-\frac{0}{1} \times \log_2 \left(\frac{0}{1}\right)\right) \\ &: 0 \\ - \text{Entropy [Pakan Ikan]} &: \left(-\frac{1}{2} \times \log_2 \left(\frac{1}{2}\right)\right) + \left(-\frac{1}{2} \times \log_2 \left(\frac{1}{2}\right)\right) \\ &: 1 \end{aligned}$$

3. *Entropy* harga

$$\begin{aligned} - \text{Entropy [Harga, murah]} &: \left(-\frac{8}{14} \times \log_2 \left(\frac{8}{14}\right)\right) + \left(-\frac{6}{14} \times \log_2 \left(\frac{6}{14}\right)\right) \\ &: 0,985228136 \\ - \text{Entropy [Harga, mahal]} &: \left(-\frac{0}{6} \times \log_2 \left(\frac{0}{6}\right)\right) + \left(-\frac{6}{6} \times \log_2 \left(\frac{6}{6}\right)\right) \\ &: 0 \end{aligned}$$

4. *Entropy* quantity

$$\begin{aligned} - \text{Entropy [Quantity, sedikit]} &: \left(-\frac{0}{12} \times \log_2 \left(\frac{0}{12}\right)\right) + \left(-\frac{12}{12} \times \log_2 \left(\frac{12}{12}\right)\right) \\ &: 0 \end{aligned}$$

$$- \text{Entropy [Quantity, banyak]} : \left(-\frac{8}{20} \times \log_2 \left(\frac{8}{20} \right) \right) + \left(-\frac{0}{20} \times \log_2 \left(\frac{0}{20} \right) \right) + \left(-\frac{12}{20} \times \log_2 \left(\frac{12}{20} \right) \right)$$

$$: 0$$

b) Mencari Nilai *Gain*

1. *Gain* Kategori Pakan

$$= 0,970950594 - \left(\left(\frac{9}{20} \right) * (0,991076059) + \left(\frac{3}{20} \right) * (0,918295834) + \left(\frac{5}{20} \right) * (0,721928094) + \left(\frac{1}{20} \right) * 0 + \left(\frac{2}{20} \right) * 1 \right)$$

$$= 0,306739968$$
2. *Gain* Harga

$$= 0,970950594 - \left(\left(\frac{14}{20} \right) * (0,985228136) + \left(\frac{6}{20} \right) * 0 \right)$$

$$= 0,281290899$$
3. *Gain* Quantity

$$= 0,970950594 - \left(\left(\frac{12}{20} \right) * 0 + \left(\frac{8}{20} \right) * 0 \right)$$

$$= 0$$

Untuk memilih atribut sebagai akar, didasarkan pada nilai *gain* yang paling tinggi dari atribut-atribut yang tersedia. Nilai *entropy* digunakan dalam menghitung nilai *gain*.

Tabel berikut menunjukkan hasil perhitungan nilai *entropy* dan *gain*:

Tabel 3. Perhitungan *Entropy* dan *Gain*

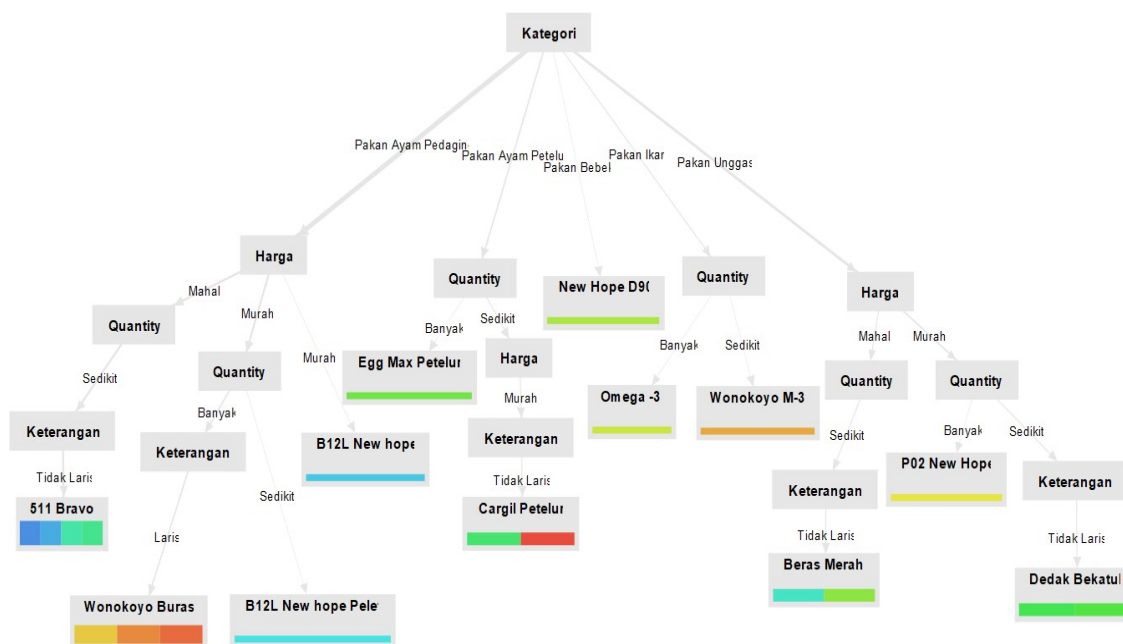
Atribut	Keterangan	Total Kasus	Laris	Kurang Laris	Entropy	Gain
Total		20	8	12	0.970950594	
Kategori Pakan						0.306739968
	Pakan Ayam Pedaging	9	4	5	0.99107606	
	Pakan Ayam Petelur	3	1	2	0.918295834	
	Pakan Unggas	5	1	4	0.721928095	
	Pakan Bebek	1	1	0	0	
	Pakan Ikan	2	1	1	1	
Harga						0.281290899
	Murah	14	8	6	0.985228136	
	Mahal	6	0	6	0	
Quantity						
	Sedikit	12	0	12	0	
	Banyak	8	8	0	0	

Sumber : Hasil Penelitian (2024)

Berdasarkan hasil perhitungan tabel, terlihat bahwa atribut kategori memiliki nilai *gain* paling tinggi, oleh karena itu, atribut kategori dapat dijadikan node akar.

4. Pengolahan Data Dengan *Decision Tree*

Berikut hasil pengolahan data menggunakan model pohon keputusan dengan menggunakan software Rapidminer, dapat dilihat pada gambar dibawah ini:



Sumber: Hasil Penelitian (2024)

Gambar 3. *Decision Tree*

Gambar 3 di atas adalah representasi grafis yang mengelompokkan berbagai jenis pakan ternak berdasarkan kategori, harga, kuantitas, dan informasi lainnya. Diagram ini bermanfaat untuk memahami karakteristik dan perbandingan antar jenis pakan. Struktur diagram ini bersifat hirarkis, dengan "Kategori" sebagai pusatnya. Terdapat empat kategori utama pakan ayam, yaitu "Pakan Ayam Pedaging", "Pakan Ayam Petelur", "Pakan Ikan", dan "Pakan Unggas". Setiap kategori memiliki sub-kategori yang mencakup nama produk pakan seperti "Wonokoyo Buras", "Egg Max Petelur", "Omega-3", dan sebagainya. Setiap produk pakan dilengkapi dengan label yang memberikan informasi tentang harga (mahal atau murah), jumlah (banyak atau sedikit), dan informasi tambahan. Informasi tambahan mencakup keterangan seperti "Tidak Laris", "Laris", "Cargil Petelur", "Beras Merah", dan lain-lain. Penggunaan warna yang berbeda digunakan untuk membedakan antara produk-produk tersebut.

KESIMPULAN DAN REKOMENDASI

Berdasarkan penelitian yang dilakukan terkait dengan penjualan pakan ternak terlaris, maka dapat diambil beberapa kesimpulan adanya data yang didapatkan dari PT Fresh X diolah dengan proses *data mining* dengan menerapkan algoritma C4.5 mulai dari penerapan *data preprocessing*, penentuan *Entropy* dan *Gain* hingga langkah terakhir yaitu penentuan pohon keputusan yang terbentuk berdasarkan hasil perhitungan *entropy* dan *gain*, dan dengan penerapan algoritma C4.5 dapat memberikan hasil yang tepat terkait dengan penjualan terlaris dan tidak laris sehingga tidak kesulitan dalam menangani permasalahan tersebut. Hasil penelitian menunjukkan *gain* tertinggi terdapat pada kategori pakan dengan nilai 0.306739968 dan *entropy* pakan ayam pedaging dengan nilai 0.99107606. Hal ini menunjukkan bahwa pakan ayam pedaging merupakan produk paling laris berdasarkan hasil pengolahan data.

REFERENSI

- Abdullah, R. W., Hartanti, D., Permatasari, H., Septyanto, A. W., & Bagaskara, Y. A. (2022). Penerapan Data Mining untuk Memprediksi Jumlah Produk Terlaris Menggunakan Algoritma Naive Bayes Studi Kasus (Toko Prapti). *Jurnal Ilmiah Informatika Global*, 13(1), 20–27. <https://doi.org/10.36982/jiig.v13i1.2060>
- Anggraini, S., Defit, S., & Nurcahyo, G. W. (2018). Analisis Data Mining Penjualan Ban Menggunakan Algoritma C4.5. *Jurnal Ilmu Teknik Elektro Komputer Dan Informatika*, 4(2), 136–143. <https://core.ac.uk/download/pdf/295348196.pdf>
- Arifin, M. F., & Fitriana, D. (2018). Penerapan Algoritma Klasifikasi C4.5 dalam Rekomendasi Penerimaan Mitra Penjualan Studi Kasus : PT Atria Artha Persada. *IncomTech*, 8(2), 87–102. <https://doi.org/10.22441/incomtech.v8i1.2198>
- Azwanti, N. (2018). Analisa Algoritma C4.5 Untuk Memprediksi Penjualan Motor Pada Pt. Capella Dinamik Nusantara Cabang Muka Kuning. *Informatika Mulawarman : Jurnal Ilmiah Ilmu Komputer*, 13(1), 33. <https://doi.org/10.30872/jim.v13i1.629>
- Cynthia, E. P., & Ismanto, E. (2018). Metode Decision Tree Algoritma C.45 Dalam Mengklasifikasi Data Penjualan Bisnis Gerai Makanan Cepat Saji. *Jurasik (Jurnal*

Riset Sistem Informasi Dan Teknik Informatika), 3(3), 1.
<https://doi.org/10.30645/jurasik.v3i0.60>

- Diansyah, T. M., & Exprada, Y. (2022). Penerapan Algoritma C4.5 Untuk Memprediksi Penjualan Lele Pada Kolam Pancing Galatama. *Jurnal Sistem Komputer Dan Informatika (JSON)*, 3(4), 567–574. <https://doi.org/10.30865/json.v3i4.4264>
- Dwiasnati, S., Gunawan, W., Oprasto, R. R., Lubis, B. O., & Santoso, B. (2023). *Algoritma dan Pemrograman Implementasi Menggunakan Python* (Y. Rahmanto (ed.); Issue 1). CV. Keranjang Teknologi Media.
- Iriadi, N., & Nuraeni, N. (2016). Kajian Penerapan Metode Klasifikasi Data Kelayakan Kredit Pada Bank Mayapada Jakarta. *Jurnal Teknik Komputer AMIK BSI*, II(1), 132–137.
- Iriadi, N., Priatno, P., & Ishaq, A. (2020). *Penerapan Data Mining dengan Rapid Miner; Konsep Data Maining, Data Warehouse, Metode, Model, Teknik*. Graha Ilmu.
- Kodari, I., Irawan, B. H., & Mustamil, M. (2020). Prediksi Laba Penjualan Menggunakan Metode Algoritma C4.5 Pada PT. Basunjaya Nastari. *EduTic*, 6(2), 61–66. <https://doi.org/10.21107/edutic.v6i2.6344>
- Lalo, A. K., Batarius, P., & Siki, Y. C. H. (2021). Implementasi Algoritma C4.5 Untuk Klasifikasi Penjualan Barang di Swalayan Dutalia. *Jurnal Teknik Informatika Unika St. Thomas (JTIUST)*, 06(1), 1–12.
- Lia, E., & Setyo Wibagso, S. (2022). Implementasi Algoritma C4.5 Untuk Penentuan Produk Terlaris (Studi Kasus: Toko XYZ Palembang). *JTSI*, 3(2), 243–255.
- Lubis, B. O. (2014). Peranan Pengetahuan Desain Komunikasi Visual dalam Pengajaran Matakuliah Interaksi Manusia dan Komputer. *Seminar Nasional Aplikasi Sains & Teknologi (SNAST)*, A-413-A-420.
- Pritalia, G. L. (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>
- Salim, A., & Lubis, B. O. (2017). Sistem Pendukung Keputusan Pemilihan Jenis Tablet Gaming dengan Menggunakan Analytical Hierarchy Process. *Prosiding Seminar*

Nasional Energi Telekomunikasi Dan Otomasi (SNETO), ISBN: 978-602-74127-4-3, 1–9.

Salim, A., Lubis, B. O., & Haidir, A. (2022). Penentuan Karyawan Terbaik Dengan Metode TOPSIS pada PT Regency Motor. *Saintekom*, 12(1), 92–102.

Suntoro, J. (2019). *Data mining : algoritma dan implementasi dengan pemrograman PHP*. Elex Media Komputindo.

Suyanto, S. (2017). *Data Mining Untuk Klasifikasi Dan Klasterisasi Data*. Informatika.