

## Klasifikasi Rasa Susu Almond Berdasarkan Minat Pelanggan Menggunakan Algoritma C4.5 di Home Made MamiNia Ice Pematangsiantar

Febriani<sup>1\*</sup>, Indra Gunawan<sup>2</sup>, Rafiq Dew<sup>3</sup>, Dedy Hartama<sup>4</sup>, Muhammad Ridwan Lubis<sup>5</sup>

<sup>1,2,3,4,5</sup>Program Studi Teknik Informatika, STIKOM Tunas Bangsa, Pematangsiantar  
Email: <sup>1\*</sup>febriani0300@gmail.com

### ABSTRACT

*Almond milk is a plant-based milk drink made from almonds. Almond juice has a paler color and a thicker texture than regular milk. As for the benefits of almond milk in the world of health, namely to increase breast milk production, prevent high blood pressure, strengthen immunity, protect bone health, maintain baby's heart health, prevent free radicals, facilitate digestion, make skin bright after childbirth. Data mining is part of the Knowledge Discovery in Database (KDD) process stage. Therefore, the authors provide a classification solution in Data Mining is done using the Algorithm C4.5. With the Algorithm C4.5 you will get a Decision Tree that is easy to understand and easy to understand. Thus, it can help the owner determine the almond milk production plan that is most in demand and without worrying about excess goods or shortages of ingredients.*

Keywords: *Almond Milk, Data Mining, Algorithm C4.5.*

### ABSTRAK

Susu almond adalah minuman susu nabati yang dibuat dari kacang almond. Sari kacang almond memiliki warna yang lebih pucat dan tekstur yang lebih kental dibandingkan susu biasanya. Adapun manfaat susu almond dalam dunia kesehatan yaitu untuk meningkatkan produksi ASI, mencegah tekanan darah tinggi, memperkuat kekebalan tubuh, melindungi kesehatan tulang, menjaga kesehatan jantung bayi, penangkal radikal bebas, memperlancar pencernaan, membuat kulit cerah setelah melahirkan. Data mining merupakan bagian dari tahapan proses Knowledge Discovery in Database (KDD). Oleh sebab itu penulis memberikan solusi klasifikasi dalam Data Mining dilakukan dengan menggunakan Algoritma C4.5. Dengan Algoritma C4.5 akan didapatkan sebuah Pohon Keputusan yang mudah dipahami dan mudah dimengerti. Dengan demikian dapat membantu pemilik menentukan perencanaan produksi susu almond yang paling banyak diminati serta tanpa khawatir terjadi kelebihan barang atau kekurangan bahan.

Kata Kunci: Susu Almond, Data Mining, Algoritma C4.5.

### 1. Pendahuluan

Minat adalah timbulnya perasaan suka terhadap suatu objek yang ditandai dengan adanya rasa senang atau tertarik terhadap sesuatu tanpa ada paksaan. Minat timbul karena seseorang merasa senang atau tertarik terhadap sesuatu disertai dengan adanya kecenderungan untuk berhubungan lebih aktif terhadap objek tersebut. "Minat pembeli adalah niat untuk melakukan pembelian pada waktu yang akan datang" [1]. Hal ini bertujuan untuk mempertahankan para pelanggan sehingga setiap pengusaha harus dapat memahami secara seksama harapan-harapan apa saja yang diinginkan semua pelanggannya, maka hal ini membuat pengusaha harus dapat mengetahui sampai mana tingkat rasa susu almond yang diminati dari setiap

pelanggan-pelanggannya. Setiap pelanggan mungkin memiliki rasa susu almond dengan minat yang berbeda, yaitu bila rasa susu almond yang dipilih tidak sesuai harapannya setelah dikonsumsi, maka pelanggan tersebut akan merasa tidak puas dan mencoba rasa susu almond yang lain. Namun bila terjadi sebaliknya, yaitu rasa susu almond sesuai dengan harapannya, maka pelanggan akan merasa puas sehingga suatu saat akan kembali mengkonsumsi atau menjadi pelanggan tetap produk susu almond tersebut sesuai dengan pilihan rasa yang disukai.

Data Mining adalah proses pengambilan pola pada data yang akan di proses sehingga *output* tersebut berupa informasi yang sangat penting dan terdapat beberapa

istilah lain yang memiliki makna yang sama[2]. Data Mining terdiri dari beberapa fase yaitu Estimasi, Clustering, Prediksi, Klasifikasi, dan Asosiasi. Data Mining memiliki salah satu metode yaitu *Decision Tree*. *Decision Tree* dengan Algoritma C4.5 merupakan klasifikasi atau segmentasi pengelompokan dan bersifat prediktif[3].

Klasifikasi merupakan salah satu proses pada data mining yang bertujuan untuk menemukan pola yang berharga dari salah satunya klasifikasi yang menggunakan refrensi struktur pohon keputusan dimana setiap *node* mempresentasikan atribut dan mempartisi kasus dalam cabang.

Contohnya pada UMKM Maminia *Ice* yang terletak di Jalan Karomah Karang Sari Pematangsiantar Sumatera Utara. Maminia *Ice* adalah sebuah usaha rumahan yang sudah berjalan sejak 2 tahun lalu. Maminia *Ice* juga memproduksi berbagai macam minuman *ice* seperti; *Almond Milk*, Es Teler, Manggo Sago, Nangka Sago, Jus Kedondong Asamboi, Jus Markijer (Markisa Jeruk), Es Kacang Hijau Kelapa Muda, dan juga menyediakan Madu Hutan asli. Disini penulis mengambil sampel untuk *Almond Milk* yang memiliki 3 varian rasa, yaitu *Almond Milk Original*, *Almond Milk Avocado*, dan *Almond Milk Chocolate*.

Algoritma C4.5 adalah salah satu algoritma yang digunakan untuk membentuk pohon keputusan [4]. Algoritma C4.5 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 [5]. Algoritma ini akan digunakan untuk menentukan rasa susu almond manakah yang paling diminati dan laku atau kurang laku dipasaran dan diharapkan dapat membantu *owner* agar kestabilan penjualan dapat berjalan. Dalam mencari klasifikasi rasa susu almond yang paling diminati oleh pelanggan, dengan metode Algoritma C4.5 penulis mengambil data tahun 2020 sebanyak 100 data pelanggan.

Berdasarkan masalah diatas maka penulis tertarik untuk meneliti Klasifikasi Rasa Susu Almond Berdasarkan Minat Pelanggan Menggunakan Algoritma C4.5 di Home Made Maminia Ice Pematangsiantar.

## 2. Metode Penelitian

### 2.1. Data Mining

Data mining merupakan proses perulangan dan memerlukan interaksi manusia dalam prosesnya untuk menemukan pola atau model baru yang dapat

Tabel 2. Data yang akan diolah pada *RapidMiner Studio Versi 5.3*

Nama	(C1)	(C2)	(C3)	Hasil
Rm. Riadho	Suka	Suka	Tidak Suka	Minat
Rm. Al Husni	Suka	Tidak Suka	Cukup Suka	Minat

digeneralisasi untuk masa yang akan datang dan bermanfaat jika digunakan untuk melakukan suatu tindakan. Ada beberapa metode yang dapat digunakan dalam data mining yaitu *Association Rule*, *Apriori*, *Roughset*, *K-Means*, *Algoritma C4.5* dan lain-lain [6].

Data mining adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer (*machine learning*) untuk menganalisis dan mengekstraksi pengetahuan (*knowledge*) secara otomatis[7].

### 2.2. Algoritma C4.5

*Algoritma C4.5* adalah salah satu metode klasifikasi dari data mining yang digunakan untuk mengkonstruksikan pohon keputusan (*decision tree*) [8]. *Algoritma C4.5* adalah algoritma yang sudah banyak dikenal dan digunakan untuk klasifikasi data yang memiliki atribut-atribut numerik dan kategorial. *Algoritma C4.5* merupakan pengembangan dari algoritma ID3[9].

Rumus menghitung *entropy* seperti rumus 1.

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

Dimana S adalah himpunan kasus, i adalah fitur, n adalah jumlah partisi S dan pi adalah proporsi dari Si terhadap S.

Rumus menghitung *gain* seperti rumus 2.

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

Dimana S adalah himpunan kasus, n adalah jumlah partisi atribut A, |si| adalah jumlah kasus pada partisi ke-i, |S| adalah jumlah kasus dalam S.

### 2.3. RapidMiner

RapidMiner merupakan platform perangkat lunak ilmu data yang dikembangkan oleh perusahaan yang bernama sama dengan pihak yang menyediakan lingkungan terintegrasi untuk persiapan data, penambahan teks, dan analisis produktif[10].

### 2.4. Analisa Data

Data yang digunakan pada penelitian ini yaitu data tahun 2020 dengan data sebanyak 100 pelanggan. Data yang digunakan sebagai penelitian dalam mengklasifikasi rasa susu almond yang paling diminati terdiri dari atribut C1 (*Almond Milk Original*), C2 (*Almond Milk Chocolate*), C3 (*Almond Milk Avocado*) dan (Hasil) dan dilakukan pengolahan data, maka hasilnya seperti tabel dibawah ini :

Nova	Cukup Suka	Tidak Suka	Cukup Suka	Tidak Minat
Herlina	Tidak Suka	Tidak Suka	Tidak Suka	Tidak Minat
Fenesa	Suka	Tidak Suka	Tidak Suka	Minat
Romauli	Tidak Suka	Tidak Suka	Suka	Minat
Annesia	Cukup Suka	Tidak Suka	Tidak Suka	Tidak Minat
Irma	Cukup Suka	Suka	Tidak Suka	Minat
Vicha	Tidak Suka	Suka	Tidak Suka	Minat
Kikie	Suka	Cukup Suka	Cukup Suka	Minat
Zea Kitchen	Suka	Suka	Suka	Minat

### 3. Hasil dan Pembahasan

#### 3.1. Menghitung Nilai Entropy dan Nilai Gain

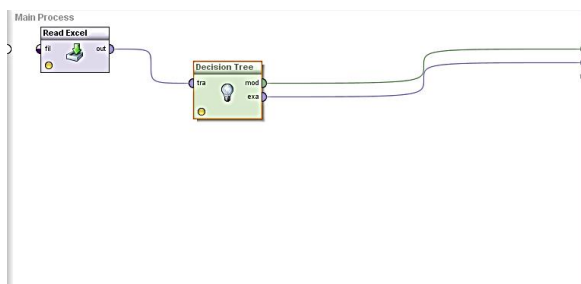
Perhitungan Nilai Entropy dan Nilai Gain menggunakan rumus dengan (Persamaan 1) dan (Persamaan 2). Sehingga menghasilkan perhitungan Node 1 yang disajikan pada Tabel 3.

Tabel 3. Perhitungan Node 1

Atribut	Jlh	M	TM	Entropy	Gain
C1	100	57	43	0,985815037	0,329345455
Suka	28	28	0	0	
Cukup Suka	18	3	15	0,650022422	
Tidak Suka	54	26	28	0,999010271	
C2					0,213616423
Suka	22	22	0	0	
Cukup Suka	20	8	12	0,970950594	
Tidak Suka	58	27	31	0,996566371	
C3					0,247299613
Suka	20	20	0	0	
Cukup Suka	20	12	8	0,970950594	
Tidak Suka	60	25	35	0,979868757	

#### 3.2. Implementasi dengan RapidMiner

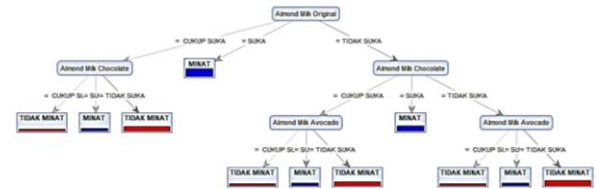
Berikut hasil proses klasifikasi dan hasil dari metode Algoritma C4.5 yang dilakukan dengan RapidMiner.



Gambar 1. Penghubung port Read excel dengan Decision Tree

Gambar 1 menjelaskan proses drag dan drop Decision Tree pada menu operators kedalam panel process. Drag dan drop kembali read excel masukkan data uji dengan cara yang sama seperti read excel pertama lalu setelah Decision Tree dan read kedua muncul maka lakukan penghubungan antara read excel dengan Decision Tree lalu ubah pada bagian criterion menjadi

information\_gain lalu klik Run maka muncul pohon keputusannya seperti pada gambar 2.



Gambar 2. Hasil Algoritma C4.5

Gambar diatas menunjukkan hasil deskripsi secara lengkap dari pohon keputusan (Decision Tree) yang telah terbentuk dengan menggunakan algoritma C4.5. Dari hasil deskripsi ini menunjukkan bahwa Almond Milk Original menjadi pohon keputusan tertinggi yang diminati pelanggan. Penggunaan Data mining Algoritma C4.5 juga baik digunakan dalam proses menggali data (data mining process) untuk menarik beberapa kesimpulan yang divisualisasikan dengan pohon keputusan (decision tree).

### 4. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, menunjukkan bahwa atribut Almond Milk Original (C1) memiliki pengaruh paling tinggi dalam klasifikasi minat tertinggi. Hasil yang diperoleh dari perhitungan manual dengan menggunakan Microsoft Excel 2010 dan hasil yang diberikan oleh Software RapidMiner Studio Versi 5.3 mendapatkan hasil berupa rule dan pohon keputusan yang sama. Dan dengan adanya penerapan data mining Algoritma C4.5 dalam melakukan klasifikasi rasa susu almond berdasarkan minat pelanggan dapat membantu owner MamiNia Ice Pematangsiantar dalam mengambil keputusan dalam menentukan rasa susu almond mana yang paling diminati pelanggannya.

#### Referensi

[1] Y. iqbal maulana and Hendri, “Memprediksi Minat Pelanggan Membeli Produk Sepatu Menggunakan Metode Decision Tree Pada PT. Zoelindo Elite Anugerah,” *Sentra Penelit. Eng. dan Edukasi*, vol. 10, no. 4, p. 70, 2018.  
 [2] A. Muzakir and R. A. Wulandari, “Model Data Mining sebagai Prediksi Penyakit Hipertensi Kehamilan dengan Teknik Decision Tree,” *Sci. J. Informatics*, vol. 3, no. 1, pp. 19–26, 2016, doi: 10.15294/sji.v3i1.4610.

- [3] A. Asroni, B. Masajeng Respati, and S. Riyadi, "Penerapan Algoritma C4.5 untuk Klasifikasi Jenis Pekerjaan Alumni di Universitas Muhammadiyah Yogyakarta," *Semesta Tek.*, vol. 21, no. 2, pp. 158–165, 2018, doi: 10.18196/st.212222.
- [4] H. Hafizan and A. N. Putri, "Penerapan Metode Klasifikasi Decision Tree Pada Status Gizi Balita Di Kabupaten Simalungun," *KESATRIA J. Penerapan Sist. Inf. (Komputer Manajemen)*, vol. 1, no. 2, pp. 68–72, 2020, doi: 10.30645/kesatria.v1i2.23.
- [5] Rismayanti, "Decision Tree Penentuan Masa Studi Mahasiswa Prodi Teknik Informatika (Studi Kasus: Fakultas Teknik dan Komputer Universitas Harapan Medan)," *J. Sist. Inf.*, vol. 02, no. 01, pp. 16–24, 2018.
- [6] E. S. Pribadi, Poningsih, and H. S. Tambunan, "Analisa Tingkat Kepuasan Masyarakat Terhadap Pelayanan Pengadilan Agama Pematangsiantar Menggunakan Algoritma," *Penerapan Kecerdasan Buatan*, vol. 2, no. 1, pp. 33–40, 2020.
- [7] L. Maulida, "PENERAPAN DATAMINING DALAM MENGELOMPOKKAN KUNJUNGAN WISATAWAN KE OBJEK WISATA UNGGULAN DI PROV . DKI JAKARTA DENGAN K-MEANS," *J. Inform. Sunan Kalijaga*, vol. 2, no. 3, pp. 167–174, 2018.
- [8] W. Rumaenda, Yuciana Wilandari, and D. Safitri, "Perbandingan Klasifikasi Penyakit Hipertensi Menggunakan Regresi Logistik Biner Dan Algoritma C4.5 (Studi Kasus Upt Puskesmas Ponjong I, Gunungkidul)," *Stud. Puskesmas, U P T*, vol. 5, no. 2, pp. 299–309, 2016.
- [9] M. F. Arifin and D. Fitriannah, "Penerapan Algoritma Klasifikasi C4.5 dalam Rekomendasi Penerimaan Mitra Penjualan Studi Kasus : PT Atria Artha Persada," *J. Telekomun. dan Komput.*, vol. 8, no. 2, pp. 88–102, 2018, doi: 10.22441/incomtech.v8i1.2198.
- [10] L. N. Rani, "Klasifikasi Nasabah Menggunakan Algoritma C4 . 5 Sebagai Dasar Pemberian Kredit," *J. INOVTEK POLBENG*, vol. 1, no. 2, pp. 127–132, 2016, doi: 10.35314/isi.v1i2.131.