

PENERAPAN ALGORITMA C.45 DALAM MENGANALISA KEPUASAN PELANGGAN PADA PENJUALAN FASHION TRIFT

Indriani Yulia Viska¹
Erlin Elisa²

¹ Mahasiswa Program Studi Sistem Informasi, Universitas Putera Batam

² dosen Program Studi Sistem Informasi, Universitas Putera Batam

email: pb191510041@upbatam.ac.id

ABSTRACT

Fashion thrift sales have become one of the fastest growing industries, especially in this digital era. This research was conducted at one of the thrift businesses in the city of Batam, while the problem with this thrift fashion shop is the lack of stock of goods that attracts customers, because the store buys sack items, likewise this store does not serve purchase orders because items in sacks are available in various types. This study aims to analyze customer satisfaction in data mining techniques to generate new knowledge using the C4.5 classification algorithm, this algorithm uses a data mining method that carries out the process of extracting information about previous decisions. This information is used as information to form a decision tree pattern. Based on all the results of the research stages that have been carried out on the Application of the Decision Tree Classification with the C4.5 Algorithm, it can be concluded that the decision tree analysis produced by the calculation of the C4.5 algorithm shows that the variables that have the highest gain or which are the main factors in determining customer satisfaction are discount variables, product quality and price

Keywords: *Algorithm C4.5, Datamining, Decision Tree, Classification, Thrift Fashion*

PENDAHULUAN

Jurnal penelitian merupakan salah satu Perkembangan dunia fashion di Indonesia saat ini sangat baik, mulai dari model busana, desain busana, gaya busana dan hal-hal lain yang diperlukan bagi masyarakat sosial. Selain itu, industri fashion merupakan salah satu industri kreatif dengan potensi ekspor produk kreatif terbesar yaitu mencapai 61,13% dari total ekspor produk kreatif. Penelitian ini dilakukan di toko svf collection trift dikota batam, kepulauan riau Toko ini telah beroperasi selama sekitar satu tahun sekarang. penjualan dilakukan

secara online di media sosial Facebook secara langsung (siaran langsung). cara yang dapat dilakukan ialah menggunakan algoritma C.45. Algoritma ini menggunakan metode data mining untuk melakukan proses penambangan informasi mengenai keputusan-keputusan sebelumnya. Informasi yang ditemukan digunakan sebagai dasar untuk membentuk pola-pola dalam bentuk pohon keputusan.

KAJIAN TEORI

2.1 *Knowledge Discovery in Database (KDD)*

Menurut (Marlina, 2018), istilah data mining dan *Knowledge Discovery in Database* (KDD) sering digunakan secara bergantian untuk menggambarkan proses penggalian informasi tersembunyi dalam database besar. Padahal, kedua istilah ini memiliki konsep yang berbeda, namun saling berkaitan. Salah satu tahapan dalam keseluruhan proses KDD adalah data mining

2.2 Data Mining

Data mining, sering disebut *knowledge discovery in database* (KDD), adalah kegiatan yang mengumpulkan dan menggunakan data historis untuk menemukan keteraturan, pola, atau hubungan dalam data dalam jumlah besar. *Database Discovery* (KDD) pada dasarnya adalah proses menemukan informasi yang berguna dari kumpulan data (Elisa, 2018)

2.3 klasifikasi

Klasifikasi adalah teknik klasifikasi yang merupakan salah satu teknik *data mining* yang paling umum digunakan dalam berbagai aplikasi. Dua jenis informasi diperlukan untuk proses klasifikasi data latihan dan data uji. Data pelatihan adalah data yang digunakan oleh algoritma *data mining* untuk mempelajari metrik klasifikasi untuk mengklasifikasikan data lain, yaitu data uji. Dua data artikel teks digunakan, dibagi menjadi data pelatihan dan data uji. (Ardi Ramdani., 2022).

2.4 Pohon keputusan

Menurut (Tangkelayuk, 2022) pohon keputusan merupakan algoritma yang banyak digunakan untuk pengambilan keputusan. Pohon keputusan adalah algoritma yang cocok untuk klasifikasi atau prediksi. Pohon keputusan adalah algoritma yang cocok untuk klasifikasi

atau prediksi. sebuah pohon yang terdiri dari beberapa bagian, yaitu root node, internal node, dan terminal node. Pohon keputusan merupakan suatu struktur yang terdiri dari simpul-simpul keputusan yang saling terhubung melalui cabang-cabang, membentang mulai dari simpul akar hingga mencapai simpul daun (akhir).

2.5 Algoritma C.45

Algoritma C4.5 yaitu sebuah algoritma yang digunakan untuk membangun *decision tree* (pengambilan keputusan). Algoritma C.45 adalah salah satu algoritma induksi pohon keputusan yaitu ID3 (*Iterative Dichotomiser 3*). ID3 dikembangkan oleh J. Ross Quinlan. Dalam prosedur algoritma ID3, input berupa sampel training, label training dan atribut. Algoritma C4.5 merupakan pengembangan dari ID3.

Beberapa pengembangan yang dilakukan pada C4.5 adalah sebagai antara lain bisa mengatasi *missing value*, bisa mengatasi *continu data*, dan pruning. Sebuah objek 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. Semakin kecil nilai *Entropy* maka akan semakin *Entropy* digunakan dalam mengekstrak suatu kelas. *Entropy* digunakan untuk mengukur ketidak aslian S. Untuk memilih atribut akar, didasarkan pada nilai gain

tertinggi dari atribut-atribut yang ada. Untuk menghitung gain digunakan rumus seperti yang tertera dalam persamaan berikut.

$$Gain(A) = Entropi(S) - \sum_i^n \frac{|S_i|}{|S|} \times Entropi(S_i)$$

Rumus 2.1 Perhitungan Gain

Dimana :

S : himpunan kasus

A : atribut

N : jumlah partisi atribut A

|S_i| : jumlah kasus pada partisi ke-i

|S| : jumlah kasus dalam S

Sementara itu, perhitungan nilai entropi dapat dilihat pada persamaan 2 berikut.

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

Rumus 2.2 Rumus Entropy

Dimana :

S : himpunan kasus

A : fitur

N : jumlah partisi S

p_i : proporsi dari S_i terhadap S

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

- a. Pilih atribut sebagai akar.
- b. Buat cabang untuk tiap-tiap nilai.
- c. Bagi kasus dalam cabang.
- d. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.(Elisa, 2018)

2.6 Weka

Weka Serangkaian alat pembelajaran mesin praktis WEKA adalah aplikasi yang diproduksi oleh *University of Waikato* di Selandia Baru. Untuk penelitian, pembelajaran dan kebutuhan aplikasi

lainnya. WEKA dapat memecahkan masalah *data mining* di dunia nyata. Khususnya dalam klasifikasi berdasarkan penerapan *machine learning*. Perangkat lunak ini dibuat dengan bahasa hirarki kelas Java dengan menggunakan metode berorientasi objek dan dapat berjalan di semua platform (Saputra, 2023)

2.7 kepuasan pelanggan

Kepuasan adalah tingkat perasaan setelah membandingkan kinerja atau hasil yang diamati dengan harapan, sejauh mana tingkat yang dirasakan dari produk dan jasa memenuhi harapan pelanggan atau pembeli. Jika aktivitas menjelaskan bahwa kualitas layanan adalah kesesuaian dan kemampuan untuk menggunakan karakteristik umum produk dan layanan yang ditawarkan untuk memenuhi kebutuhan dan harapan yang diinginkan konsumen dengan karakteristik atau faktor seperti: tampilan langsung, perhatian pribadi karyawan kepada konsumen, daya tanggap, kehandalan dan jaminan. Jika kinerja memenuhi harapan, pelanggan puas. Harapan pelanggan dapat dibentuk oleh pengalaman masa lalu, komentar kerabat, serta janji dan informasi dari pemasar dan pesaing (Riyani, 2021)

2.8 Thrift

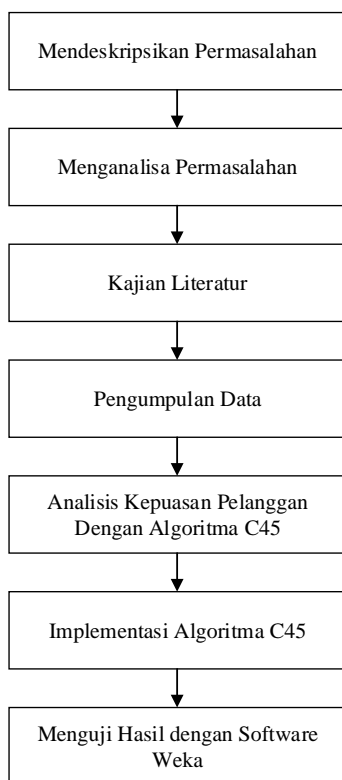
Salah satu bisnis yang paling aktif di masyarakat, terutama di kalangan anak muda adalah bisnis hemat. Kata thrift sendiri berasal dari kata thrive yang memiliki arti berkembang atau maju, sedangkan Bahasa thrifty sendiri memiliki makna cara memanfaatkan uang ataupun barang yang lainnya dengan baik dan efisien. Thrifting bisa juga diartikan yakni aktivitas membeli barang bekas, namun aktivitas ini bukan saja membeli barang bekas, akan tetapi juga berarti bagaimana kepuasan pribadi terhadap barang yang bagus dengan budget yang lebih murah

atau harga miring. mulai dari baju import hingga baju, topi, sepatu dan celana. Namun ia menjual barang dengan beberapa merk ternama seperti *Nike*, *Adidas*, *Dickies*, dll. Meski dalam beberapa tahun terakhir menjadi pemberitaan, pemerintah telah melarang pembelian dan penjualan pakaian bekas impor dengan berbagai alasan, salah satunya adalah kepedulian pemerintah terhadap penampilan yang sehat.

METODE PENELITIAN

3.1 Desain penelitian

Desain Penelitian yang akan digunakan dalam penelitian ini dirancang dan dipaparkan pada gambar dibawah ini, yakni:



Sumber : Peneliti 2023

Gambar Desain Penelitian

Berikut pemaparan dari desain penelitian yang di lakukan pada penelitian ini ialah:

1. Mendeskripsikan Permasalahan
Pada tahap ini peneliti melakukan observasi lapangan untuk mengetahui permasalahan yang ada, kemudian masalah tersebut akan di deskripsikan yang berkaitan dengan kepuasan pelanggan.
2. Menganalisa Permasalahan
Pada tahap ini peneliti akan melakukan analisa dari permasalahan yang telah ditemukan yang akan digunakan dalam penelitian ini
3. Kajian Literatur
Pada tahap ini peneliti akan menggunakan beberapa sumber pendukung yang berkaitan dengan penelitian ini
4. Pengumpulan Data
Proses ini didapatkan dari hasil observasi dan juga dari hasil tanya jawab kepada pemilik toko *svf collection thrift*
5. Analisis Kepuasan Pelanggan Dengan Algoritma C45
Selama fase ini, penggunaan jenis algoritme dari C4.5 dimulai dengan proses pembuatan pohon keputusan untuk membuat penilaian kepuasan pelanggan secara teratur.
6. Implementasi Algoritma C45
Pada tahap ini peneliti Memanfaatkan data untuk mengetahui validitas dari akar pohon pemilihan, menentukan nilai keuntungan maksimum yang dihasilkan dari setiap aspek dalam data yang disediakan, dan kemudian memutuskan nilai apa yang

- diperoleh dalam proses membangun desain matriks pilihan
- Menguji Hasil dengan Software Weka
Hasil dari perhitungan manual kemudian di Uji kan kedalam software weka untuk membuat pohon keputusan

HASIL DAN PEMBAHASAN (Arial 4.1
hasil

4.1.1 Proses algoritma C.45
Berdasarkan hasil kuisisioner yang telah dkumpulkan diambil 140 sampel untuk diolah. Maka hasil perhitungan Node1 sebagai berikut.

Tabel 1. Hasil Perhitungan *Node 1*

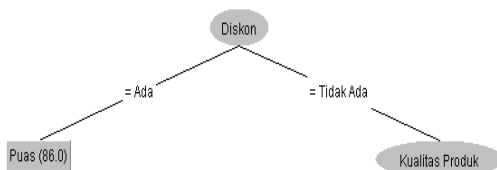
		Jumlah (s)	Puas (si)	Tidak Puas (si)	Entropy	Gain
Total		140	90	50	0.940285959	
Harga	Mu	84	72	12	0.591672779	0.222911005
	Ma	56	18	38	0.905928216	
Pelayanan	R	96	86	10	0.482066148	0.747853939
	TR	44	4	40	0.439496987	
Merk	L	76	64	12	0.629249224	0.16243055
	I	38	18	20	0.998000884	
	TAM	26	8	18	0.89049164	
Kualitas Produk	B	92	84	8	0.426228657	0.846557793
	TB	48	6	42	0.543564443	
Diskon	A	86	86	0	0	1.087222499
	TA	54	4	50	0.380946586	

(Sumber: Data Penelitian, 2019)

Variabel atribut diskon memiliki *Gain* tertinggi dibanding variabel atribut lainnya dengan nilai *Gain* 1,0872, maka, variabel atribut harga menjadi *root Node* pada pohon keputusan sementara.

Gambar 4.1. Pohon keputusan *Node 1*

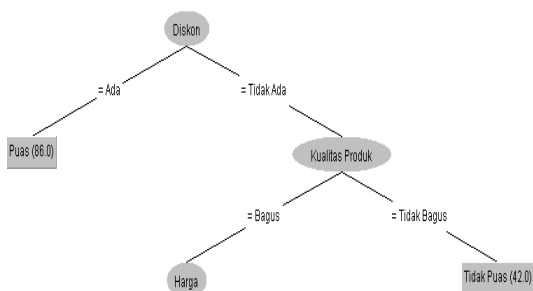
Selanjutnya akan di lakukan perhitungan nilai entrophy dan gain untuk node 2, Berikut tabel hasil akhir dari perhitungan *Node 2*:



Tabel 2. Perhitungan hasil *Node 2*

		Jumlah (s)	Puas (si)	Tidak Puas (si)	Entropy	Gain
Total A		54	4	50	0.380946586	
Harga	MU	16	4	12	0.811278124	0.140567882
	MA	38	0	38	0	
Pelayanan	R	14	4	10	0.863120569	0.157174586
	TR	40	0	40	0	
Merk	L	16	4	12	0.811278124	0.140567882
	I	20	0	20	0	
	TAM	18	0	18	0	
Kualitas Produk	B	12	4	8	0.918295834	0.176880845
	TB	42	0	42	0	

Dari perhitungan node 2 atribut harga dan merk mempunyai nilai *Gain* yang sama, tetapi kualitas produk tetap menjadi nilai *gain* tertinggi, berikut pohon keputusannya, yaitu:



Gambar 2. Pohon Keputusan *Node 2*

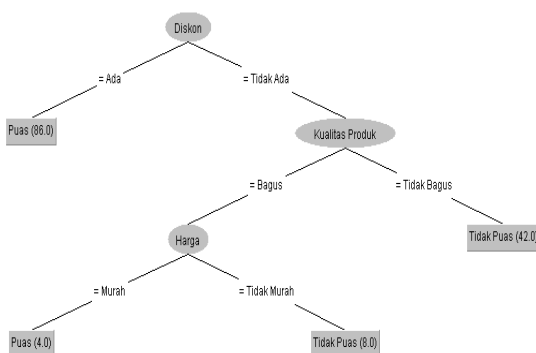
Nilai atribut Tidak bagus menjadi akhir karena tidak dapat diturunkan. Nilai entropy ditunjukkan dengan 0, tetapi nilai atribut bagus masih dapat diturunkan. Proses berikut digunakan untuk

menghitung turunan dari node 2 ke node 3 dengan cara yang sama seperti node 2 di atas. Dalam hal ini, entropy dan keuntungan dari atribut yang tersisa dihitung. Kualitas produk dengan klasifikasi bagus yang berjumlah 12 dengan 4 puas dan 8 tidak puas yaitu harga, pelayanan, dan merk. langkah selanjutnya adalah menempatkannya di tabel node 3:

Tabel 3.Perhitungan hasil *Node 3*

		Jumlah (s)	Puas (si)	Tidak Puas (si)	Entropy	Gain
Total B		12	4	8	0.918295834	
Harga	MU	4	0	4	0	0.918295834
	MA	8	0	8	0	
Pelayanan	R	4	4	0	0	0.918295834
	TR	8	0	8	0	
Merk	L	4	4	0	0	0.918295834
	I	4	0	4	0	
	TAM	4	0	4	0	

Dapat terlihat bahwa nilai gain untuk setiap atribut adalah sama, 0,9182. Oleh karena itu data pertama atau harga adalah turunan akhir dari pohon keputusan.



Gambar 3 *Node 3*

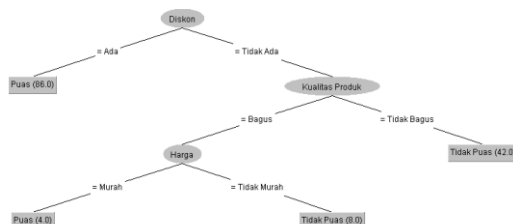
Berlandaskan dari perhitungan diatas yang sudah memiliki perhitungan sama dan sudah tidak memiliki atribut akar maka perhitungan diselesaikan.

4.1.2 Hasil Rules Pohon Keputusan

Berdasarkan pohon keputusan diatas bisat dihasilkan bahwa satu atau lebih aturan yang dihasilkan oleh pohon keputusan akhir adalah:

1. Jika diskon ada maka konsumen merasa puas
2. Jika diskon tidak ada dan kualitas produk tidak bagus, maka konsumen tidak puas
3. Jika diskon tidak ada dan kualitas produk bagus maka konsumen merasa puas
4. Jika diskon tidak ada, kualitas produk bagus dan harga murah maka konsumen merasa puas
5. Jika diskon tidak ada, kualitas produk bagus dan harga mahal maka konsumen tidak puas
6. Jika diskon ada dan kualitas produk bagus dan harga mahal maka konsumen tidak puas
7. Jika kualitas produk tidak bagus konsumen tidak puas

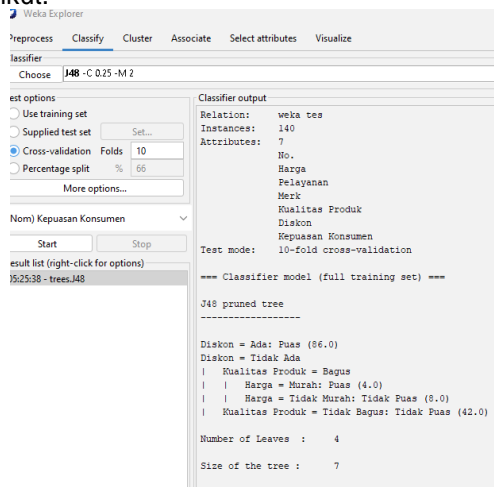
8. Jika kualitas produk bagus, harga mahal maka konsumen tidak puas
9. Jika kualitas produk bagus, harga murah maka konsumen merasa puas
10. Jika harga murah maka konsumen merasa puas
11. Jika harga mahal maka konsumen tidak puas



Gambar 5. Hasil *Visualize Tree* pada WEKA

4.2 Pembahasan

Untuk menguji keabsahan hasil maka penulis akan menggunakan software weka untuk perbandingan hasil, adapun hasil klasifikasi outputnya sebagai berikut.



Gambar 4. *Classify Output*

Sehingga pohon keputusan yang dibentuk sama dengan perhitungan manual dimana atribut Diskon, kualitas dan harga memiliki nilai gain yang tertinggi dan menjadikan faktor kepuasan pelanggan terhadap objek penelitian. Pohon keputusan weka dapat dilihat sebagai berikut.

SIMPULAN

1. penggunaan teknik data mining dengan metode klasifikasi algoritma C4.5 pada analisis kepuasan konsumen di Toko Svf Collection Thrift menggunakan data kuesioner dapat menghasilkan aturan keputusan atau rules berdasarkan variabel Diskon, Kualitas, dan Harga. Aturan-aturan ini dapat digunakan sebagai acuan untuk memahami faktor-faktor yang memengaruhi kepuasan konsumen di toko tersebut.
2. Hasil dari analisis pohon keputusan (decision tree) yang menggunakan algoritma C4.5 menunjukkan bahwa variabel dengan gain tertinggi, atau dengan kata lain, variabel yang menjadi faktor utama dalam menentukan kepuasan konsumen adalah diskon, kualitas produk, dan harga.

DAFTAR PUSTAKA

Ardi Ramdani, Christian Dwi Sofyan, Fauzi Ramdani, Muhamad Fauzi Arya Tama, & Muhammad Angga Rachmatsyah. (2022). Algoritma Klasifikasi Data Mining Untuk Memprediksi Masyarakat Dalam Menerima Bantuan Sosial. *Jurnal Ilmiah Sistem Informasi*, 1(2), 39–47.

<https://doi.org/10.51903/juisi.v1i2.363>

Elisa, E. (2018). Prediksi Profit Pada Perusahaan Dengan Klasifikasi Algoritma C4.5. *Klik - Kumpulan Jurnal Ilmu Komputer*, 5(2), 179. <https://doi.org/10.20527/klik.v5i2.153>

Marlina, D., Putri, N. F., Fernando, A., & Ramadhan, A. (2018). *Implementasi Algoritma K-Medoids dan K-Means untuk Pengelompokan Wilayah Sebaran Cacat pada Anak*. 4(2), 64–71.

Riyani, D., Larashat, I., & Juhana, D. (2021). Pengaruh Harga Dan Kualitas Pelayanan Terhadap Kepuasan Pelanggan. *Majalah Bisnis & IPTEK*, 14(2), 94–101. <https://doi.org/10.55208/bistek.v14i2.233>

Saputra, R., Wijaya, D., Wijaya, T., & ... (2023). Penerapan Data Mining dalam Proses Penerimaan Karyawan Baru pada Perusahaan Outsourcing (SDM) Dengan Metode C4. 5 Menggunakan Aplikasi Weka. *Jurnal Bidang ...*, 89–98.

Tangkelayuk, A. (2022). The Klasifikasi Kualitas Air Menggunakan Metode

KNN, Naïve Bayes, dan Decision Tree. *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, 9(2), 1109–1119. <https://doi.org/10.35957/jatisi.v9i2.2048>

	<p>BIODATA Penulis pertama, Indriani Yulia Viska merupakan seorang mahasiswa prodi Sistem informasi universitas Putera Batam.</p>
	<p>BIODATA Penulis kedua, Erlin Elisa, Merupakan dosen prodi Sistem informasi di Universitas Putera Batam, Seorang tenaga pendidik yang menyelesaikan pendidikan strata-1 dari Universitas Putra Indonesia YPTK Padang pada tahun 2008 dan melanjutkan pendidikan master nya di Universitas Putra Indonesia YPTK Padang pada tahun 2015.</p>