

Jurnal 7

by Plag Turnitin

Submission date: 19-Mar-2024 12:00PM (UTC+0700)

Submission ID: 2284259357

File name: Jurnal_7.pdf (242.24K)

Word count: 2513

Character count: 14622

Perhitungan Matematis Algoritma C4.5 Untuk Memprediksi Penjualan *Wallpaper* (Studi Kasus : Toko Ali Banyuanyar Probolinggo)

Nurhidayati¹, Imam Marzuki²

¹ e-mail: yati150690@gmail.com, ² imam@upm.ac.id

² Program Studi Tadris Matematika, STAI Muhammadiyah Probolinggo
² Program Studi Teknik Elektro Fakultas Teknik Universitas Panca Marga

Abstract— Coloring the walls of the house with wallpaper can indeed add a beautiful impression to a room, so that the room looks more expressive. Various motifs, colors and patterns of wallpaper can be selected. To increase wallpaper sales, sellers must be more careful about which wallpaper will be in high demand. Need recommendations for the type of wallpaper. The data required is data on the best wallpaper brands, colors, patterns, quality of materials, sizes, and prices. Then the data is classified. The method used for classification is data mining with the C4.5 algorithm. From the calculation process of the C4.5 algorithm, several rules will be obtained. This rule is used as the basis for predicting wallpaper sales.

Keywords— *Wallpaper, Mathematical Calculations, C4.5 Algorithm*

Intisari — Mewarnai dinding rumah dengan *wallpaper* memang bisa menambah kesan indah pada suatu ruangan, sehingga ruangan terlihat lebih ekspresif. Berbagai motif, warna, dan corak wallpaper dapat dipilih. Untuk meningkatkan penjualan *wallpaper*, Penjual harus lebih cermat menyediakan wallpaper mana yang akan banyak peminatnya. Perlu perekomendasi jenis jenis *wallpaper*. Data yang diperlukan yaitu data merk *wallpaper* terbaik, warna, motif, kualitas bahan, ukuran, dan harga. Kemudian data tersebut diklasifikasikan. Metode yang digunakan untuk klasifikasi adalah data mining dengan algoritma C4.5. Dari proses perhitungan algoritma C4.5 akan didapatkan beberapa *rule*. *Rule* tersebut dijadikan dasar untuk memprediksi penjualan *wallpaper*.

Kata Kunci— *Wallpaper, Perhitungan Matematis, Algoritma C.45*

I. PENDAHULUAN

Untuk memperindah dinding tidak selalu menggunakan cat dinding. Namun juga ada alternatif lain yang bisa digunakan. Salah satunya adalah *wallpaper*. Wallpaper dinding rumah atau kertas dinding merupakan hiasan dinding dengan bermacam motif dan warna. Kertas dinding digunakan untuk mengubah penampilan suatu ruang agar semakin cantik dan memiliki nilai tambah.

Toko Ali di Banyuanyar Probolinggo merupakan salah satu toko yang melayani penjualan wallpaper. Dalam kesehariannya, Toko Ali menyediakan berbagai macam motif, warna, merek, kualitas bahan, ukuran dan harga yang berbeda. Dalam menyediakan wallpaper toko ali masih memesan wallpaper dari produsen. Pemesanan wallpaper masih bersifat perkiraan terhadap selera konsumen. Kemudian toko ali memberikan rekomendasi jenis-jenis wallpaper kepada konsumen. Dan terkadang rekomendasi ini tidak sesuai dengan yang diharapkan. Konsumen ternyata tidak memilih jenis wallpaper yang disediakan toko ali. Hal ini menyebabkan pembelian wallpaper menjadi terhambat waktu karena toko ali harus memesan kembali ke produsen dalam beberapa hari sesuai dengan permintaan konsumen.

Untuk mempermudah toko ali memberikan rekomendasi yang sesuai memilih wallpaper mana yang banyak diminati konsumen sehingga jenis-jenis wallpaper akan dengan mudah disediakan, maka perlu diprediksi untuk penjualan wallpaper terbanyak.

Telah banyak penelitian yang membahas tentang prediksi penjualan wallpaper, namun dari sekian penelitian tersebut tidak menggambarkan perhitungan matematis yang mendetail untuk menjawab masalah prediksi penjualan wallpaper. Oleh karena itu dalam penelitian ini difokuskan pada perhitungan matematis penggunaan algoritma C4.5.

Penggunaan Algoritma C4.5 dipilih karena proses klasifikasi data sederhana dan cepat. Hasilnya yaitu berupa pohon keputusan yang lebih mudah dipahami dan diinterpretasikan [3]. Selanjutnya pohon keputusan dapat menghasilkan beberapa *rule* yang dapat dijadikan acuan untuk memprediksi penjualan *wallpaper* dimasa-masa yang akan datang.

II. TINJAUAN PUSTAKA

1 Algoritma C4.5 adalah algoritma yang digunakan untuk menghasilkan sebuah pohon keputusan yang dikembangkan oleh Ross Quinlan[2]. Ide dasar dari algoritma ini adalah pembuatan pohon keputusan berdasarkan pemilihan atribut yang memiliki prioritas tertinggi atau dapat disebut memiliki nilai gain tertinggi berdasarkan nilai entropy atribut tersebut sebagai poros atribut klasifikasi[8]. Pada tahapannya algoritma C4.5 memiliki 2 prinsip kerja[9], yaitu: Membuat pohon keputusan, dan membuat rule. Aturan-aturan yang terbentuk dari pohon keputusan akan membentuk suatu kondisi dalam bentuk *if then*.

Terdapat empat langkah dalam proses pembuatan pohon keputusan pada algoritma C4.5, yaitu:

1. Memilih atribut sebagai akar, didasarkan pada nilai gain tertinggi dari atribut-atribut yang ada.
2. Membuat cabang untuk masing-masing nilai, artinya membuat cabang sesuai dengan jumlah nilai variabel gain tertinggi.
3. Membagi setiap kasus dalam cabang, berdasarkan perhitungan nilai gain tertinggi dan perhitungan dilakukan setelah perhitungan nilai gain tertinggi awal dan kemudian dilakukan proses perhitungan gain tertinggi kembali tanpa meyeratakan nilai variabel gain awal.
4. Mengulangi proses dalam setiap cabang sehingga semua kasus dalam cabang memiliki kelas yang sama, mengulangi semua proses perhitungan gain tertinggi untuk masing-masing cabang kasus sampai tidak bisa lagi dilakukan proses perhitungan.

Algoritma C4.5 secara rekursif mengunjungi setiap simpul keputusan, memilih pembagian yang optimal, sampai tidak bisa dibagi lagi. Dari ketiga peneliti yang telah dilakukan tersebut, klasifikasi dengan Algoritma C4.5 digunakan oleh para peneliti sebagai solusi untuk mengambil keputusan yang diharapkan mampu membantu dalam pengambilan keputusan dengan lebih mudah dan cepat[2]. Dalam penerapan dan penggunaan algoritma C4.5, dapat digunakan untuk melakukan prediksi dan klasifikasi terhadap calon pegawai yang berpotensi untuk masuk ke dalam perusahaan dengan cara membuat pohon keputusan berdasarkan data-data yang sudah ada dan melakukan prediksi terhadap calon pegawai baru yang ingin masuk ke perusahaan[1]. Selain itu algoritma ini digunakan untuk klasifikasi predikat keberhasilan mahasiswa di sebuah universitas. Variabel yang memiliki prioritas utama terhadap predikat keberhasilan mahasiswa adalah mahasiswa yang memilih sesi perkuliahan pada pagi hari[11].

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* masing-masing atribut. Penghitungan nilai Entropy digunakan rumus seperti dalam Persamaan (1)[2].

$$Entropy(S) = \sum_{i=0}^n - p_i * \log_2(p_i) \quad (1)$$

Rumus (1) merupakan rumus yang digunakan dalam perhitungan entropy yang digunakan untuk menentukan seberapa informatif atribut tersebut. Berikut keterangannya :

- S : Himpunan kasus
- n : Jumlah partisi
- p_i : Jumlah kasus pada partisi ke-i

Information Gain adalah informasi yang didapatkan dari perubahan entropy pada suatu kumpulan data, baik melalui observasi atau bisa juga disimpulkan dengan cara melakukan partisipasi terhadap suatu set data[8].

$$GAIN(S, A) = Entropy(S) - \left| \frac{S_i}{S} \right| * Entropy(S) \quad (2)$$

Rumus (2) merupakan rumus yang digunakan dalam perhitungan *information gain* setelah melakukan perhitungan entropy. Berikut keterangannya :

- S : Himpunan kasus
- n : Jumlah partisi atribut A
- $|S_i|$: Jumlah kasus pada partisi ke-i
- |S| : Jumlah kasus dalam s

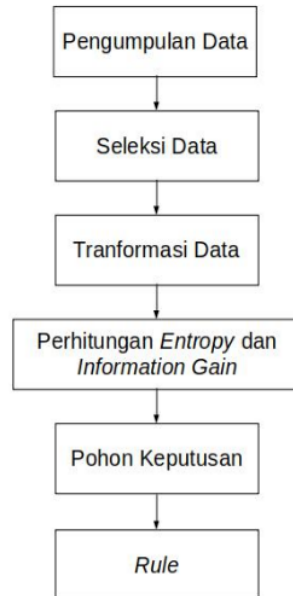
Dengan mengetahui rumus-rumus diatas, data yang telah diperoleh dapat dimasukkan dan diproses dengan algoritma C4.5 untuk proses pembuatan pohon keputusan.

Pohon keputusan adalah sebuah struktur yang dapat digunakan untuk mengubah data menjadi pohon keputusan yang akan menghasilkan aturan-aturan keputusan besar menjadi himpunan-himpunan record yang lebih kecil dengan menerapkan

serangkaian aturan keputusan[2]. Pohon keputusan yang dihasilkan oleh algoritma C4.5 dapat digunakan untuk klasifikasi[3]. Pohon keputusan adalah salah satu teknik penambangan data yang paling populer untuk penemuan pengetahuan. Secara sistematis menganalisis dan mengekstrak aturan untuk tujuan klasifikasi / prediksi[1].

III. METODE

Metode Penelitian yang digunakan dalam penelitian ini, menggunakan rancangan penelitian yang ditunjukkan pada Gambar 1.



Gambar 1. Blok diagram penelitian

Berdasarkan blok diagram pada gambar 1 dijelaskan sebagai berikut :

Pengumpulan Data

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

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[16].

Transformasi Data

Transformasi data adalah proses mentransformasi atau mengubah data ke dalam bentuk yang sesuai, agar dapat di proses dengan perhitungan algoritma C4.5. Transformasi Data adalah proses mentransformasi atau menggabungkan data ke dalam bentuk yang sesuai untuk penggalan lewat operasi summary atau aggregation[16].

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.

Pohon Keputusan

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.

Rule

Rule adalah uraian penjelasan yang merepresentasikan sebuah pohon keputusan.

IV. HASIL DAN PEMBAHASAN

Data penjualan wallpaper toko ali dapat dituangkan dalam tabel 1 berikut.

Tabel 1. Penjualan wallpaper di toko ali

No	Tekstur	Motif	Kualitas	Harga	Peminat
1	Timbul	Tradisional	Tinggi	MURAH	Sedikit
2	Timbul	Tradisional	Tinggi	MAHAL	Sedikit
3	Biasa	Tradisional	Tinggi	MURAH	Banyak
4	Anti Air	Modem	Tinggi	MURAH	Banyak
5	Anti Air	Dunia Anak	Rendah	MURAH	Banyak
6	Anti Air	Dunia Anak	Rendah	MAHAL	Sedikit
7	Biasa	Dunia Anak	Rendah	MAHAL	Banyak
8	Timbul	Modem	Tinggi	MURAH	Sedikit
9	Timbul	Dunia Anak	Rendah	MURAH	Banyak
10	Anti Air	Modem	Rendah	MURAH	Banyak
11	Timbul	Modem	Tinggi	MAHAL	Banyak
12	Biasa	Modem	Tinggi	MAHAL	Banyak
13	Biasa	Tradisional	Rendah	MURAH	Banyak
14	Anti Air	Modem	Tinggi	MAHAL	Sedikit

Pada kasus yang terdapat pada tabel 1 akan dibuat pohon keputusan untuk menentukan banyak atau sedikitnya peminat dengan melihat tekstur, motif, kualitas, dan harga.

Langkah-langkah penyelesaian kasus pada table 2.1 adalah:

- Menghitung jumlah kasus, jumlah kasus untuk keputusan Yes, jumlah kasus untuk keputusan No, dan Entropy dari semua kasus dan kasus yang dibagi berdasarkan atribut TEKSTUR, MOTIF, KUALITAS, dan HARGA.
- Lakukan perhitungan Gain untuk setiap atribut. Hasil perhitungan ditunjukkan oleh tabel 2.

Tabel 2. Perhitungan Node 1

Node	Atribut	Nilai	Jumlah Kasus (S)	Sedikit (S1)	Banyak (S2)	Entropy	Gain
1	TOTAL		14	4	10	0.8631205	
	TEKSTUR						0.2585210
		Timbul	4	0	4		
		Biasa	5	1	4	0.721928	
		Anti Air	5	3	2	0.970950	
	MOTIF						0.1838509
		Tradisional	4	0	4	0	
		Modern	4	2	2	1	
		Dunia Anak	6	2	4	0.9182958	
	KUALITAS						0.3705065
		Tinggi	7	4	3	0.9852281	
		Rendah	7	0	7	0	
	HARGA						0.0059777
		Mahal	8	2	6	0.8112781	
		Murah	6	4	2	0.9182958	

Baris total kolom Entropy pada table 2 dihitung dengan rumus (1) :

$$Entropy(Total) = \left(-\frac{4}{14} * \log_2 \left(\frac{4}{14}\right)\right) + \left(-\frac{10}{14} * \log_2 \left(\frac{10}{14}\right)\right) \\ = 0.863120569$$

Nilai *Information Gain* pada baris TEKSTUR dihitung dengan rumus (2) :

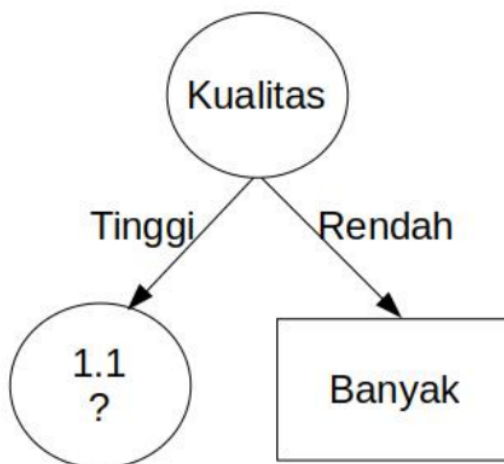
$$Gain(Total, Tekstur) \\ = Entropy(Total) - \sum_{i=1}^n \frac{|Tekstur|}{|Total|} * Entropy(Tekstur)$$

$$= 0.863120569 - \left(\left(\frac{4}{14} * 0 \right) + \left(\frac{5}{14} * 0.723 \right) + \left(\frac{5}{14} * 0.97 \right) \right)$$

$$= 0.23$$

Dari tabel 2 dapat diketahui bahwa atribut dengan *Information Gain* tertinggi adalah KUALITAS, yaitu sebesar 0.37. Jadi KUALITAS yang menjadi node akar. Ada dua nilai atribut dari KUALITAS, yaitu TINGGI dan RENDAH. Dari kedua atribut, nilai atribut NORMAL adalah 1, yaitu keputusan Yes, sehingga tidak perlu perhitungan lebih lanjut, tetapi untuk HIGH masih perlu dilakukan perhitungan lagi.

Dari hasil tersebut dapat digambarkan pohon keputusan sementara seperti gambar 2.



Gambar 2. Pohon keputusan hasil perhitungan Node 1

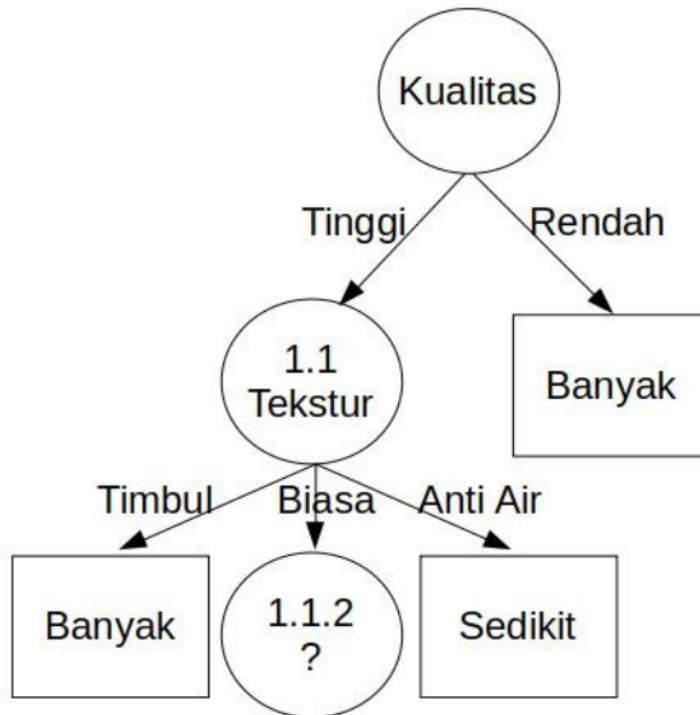
Menghitung jumlah kasus, jumlah kasus untuk keputusan BANYAK, jumlah kasus untuk keputusan SEDIKIT, dan ENTROPY dari semua kasus dan kasus yang dibagi berdasarkan atribut TEKSTUR, MOTIF, dan HARGA yang dapat menjadi node akar dari nilai atribut TINGGI. Kemudian lakukan perhitungan *Information Gain* untuk tiap-tiap atribut. Hasil perhitungan seperti yang terdapat pada tabel 3 :

Tabel 2. Perhitungan Node 1.1

Node	Atribut	Nilai	Jumlah Kasus (S)	Sedikit (S1)	Banyak (S2)	Entropy	Gain
1	TOTAL		7	4	3	0.9852281	
	TEKSTUR						0.69951385
		Timbul	2	0	2	0	
		Biasa	2	1	1	1	
		Anti Air	3	3	0	0	
	MOTIF						0.02024420
		Tradisional	0	0	0	0	
		Modern	4	2	1	0.9182958	
		Dunia Anak	3	2	2	1	
	HARGA						0.02024420
		Mahal	4	2	2	1	
		Murah	3	2	1	0.9182958	

Dari hasil tabel 3 dapat diketahui bahwa atribut dengan *Information Gain* tertinggi adalah TEKSTUR, yaitu sebesar 0.67. Jadi TEKSTUR dapat menjadi node cabang dari nilai atribut TINGGI. Ada tiga nilai atribut dari TEKSTUR, yaitu TIMBUL, BIASA dan ANTI AIR. Dari ketiga nilai atribut tersebut, nilai atribut TIMBUL adalah 2, yaitu keputusannya BANYAK dan nilai atribut ANTI AIR adalah 0 sehingga menjadi keputusan SEDIKIT. Oleh Karena itu atribut ANTI AIR tidak perlu dilakukan perhitungan lebih lanjut, tetapi nilai atribut BIASA masih perlu dilakukan perhitungan lagi.

Pohon keputusan yang terbentuk sampai tahap ini adalah terlihat pada gambar 3.



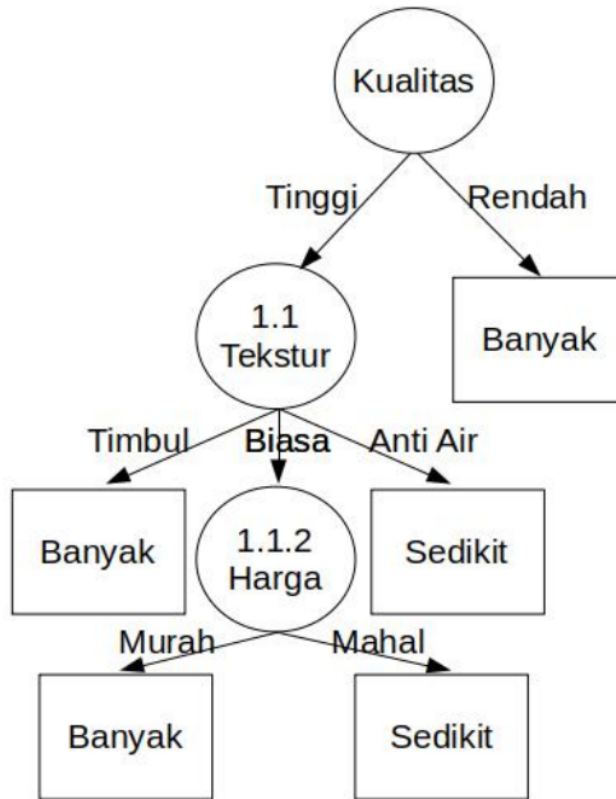
Gambar 3. Pohon keputusan hasil perhitungan Node 1.1

Menghitung jumlah kasus, jumlah kasus untuk keputusan Yes, jumlah kasus untuk keputusan No, dan *Entropy* dari semua kasus dan kasus yang dibagi berdasarkan atribut MOTIF dan HARGA yang dapat menjadi node cabang dari nilai atribut BIASA. Kemudian lakukan perhitungan *Information Gain* untuk tiap-tiap atribut. Hasil perhitungan seperti yang terdapat pada tabel 3 :

Tabel 3. Perhitungan Node 1.1.2

Node	Atribut	Nilai	Jumlah Kasus (S)	Sedikit (S1)	Banyak (S2)	Entropy	Gain
1	TOTAL		2	1	1	1	
	MOTIF						0
		Tradisional	0	0	0	0	
		Modern	0	0	0	0	
		Dunia Anak	2	1	1	1	
	HARGA						1
		Mahal	1	0	1	0	
		Murah	1	1	0	0	

Dari hasil tabel 3 dapat diketahui bahwa atribut dengan *Information Gain* tertinggi adalah HARGA, yaitu sebesar 1. Jadi HARGA dapat menjadi node cabang dari nilai atribut BIASA. Ada dua nilai atribut dari HARGA, yaitu MAHAL dan MURAH. Dari kedua nilai atribut tersebut, nilai atribut MAHAL adalah 1, yaitu keputusannya BANYAK dan nilai atribut MURAH menjadi keputusan SEDIKIT, sehingga tidak perlu dilakukan perhitungan lebih lanjut untuk nilai atribut ini. Pohon keputusan yang terbentuk sampai tahap ini adalah terlihat pada gambar 4 berikut:



Gambar 4. Pohon keputusan hasil perhitungan Node 1.1.1

Gambar 4 merupakan gambar akhir dari proses klasifikasi menggunakan algoritma C4.5 pada kasus penjualan wallpaper. Berdasarkan gambar di atas dapat diperoleh rule :

- If KUALITAS RENDAH maka PEMINAT adalah BANYAK
- If KUALITAS TINGGI and TEKSTUR TIMBUL maka PEMINAT adalah BANYAK
- If KUALITAS TINGGI dan TEKSTUR BIASA dan HARGA MURAH maka PEMINAT BANYAK
- If KUALITAS TINGGI dan TEKSTUR BIASA dan HARGA MAHAL maka PEMINAT SEDIKIT
- If KUALITAS TINGGI dan TEKSTUR Terakhit jika humadity adalah high, outlook adalah ANTI AIR, maka PEMINAT SEDIKIT

2 V. KESIMPULAN

Berdasarkan hasil penelitian yang dilakukan, maka penulis dapat menarik kesimpulan dalam penjualan wallpaper dengan menggunakan metode klasifikasi khususnya Algoritma C4.5 akan bermanfaat sekali dalam Proses pengambilan keputusan dalam penjualan wallpaper.

1. Yang berpengaruh dalam penjualan wallpaper adalah faktor kualitas, tekstur dan harga.
2. Hasil akhir dari proses perhitungan algoritma adalah pohon keputusan yang menghasilkan beberapa rule yang dapat dimanfaatkan untuk memprediksi penjualan wallpaper dimasa yang akan datang.

REFERENSI

- [1] Fricles A S. 2018. Analisa Decision Tree Dalam Pengolahan Data Siswa. STIMIK Pelita Nusantara. Medan.
- [2] Rina N. 2016. Teknik Data Mining : Algoritma C 4.5. Ilmukomputer.com
- [3] Siti N K. 2018. Penerapan Algoritma C4.5 Untuk Penentuan Kelayakan Kredit. STIMIK Nusa Mandiri Jakarta. Jakarta.
- [4] Muhammad F A, Devi F. Penerapan Algoritma Klasifikasi C4.5 dalam Rekomendasi Penerimaan Mitra Penjualan Studi Kasus : PT Atria Artha Persada. Universitas Mercubuana Jakarta. Jakarta.

Jurnal 7

ORIGINALITY REPORT

15%

SIMILARITY INDEX

9%

INTERNET SOURCES

18%

PUBLICATIONS

29%

STUDENT PAPERS

PRIMARY SOURCES

1

eprints.uty.ac.id

Internet Source

9%

2

Imam Marzuki, Tri Prihatiningsih. "Prediksi Kelayakan Ekonomi Pada Pengolahan Limbah Cair Tahu di CV. Proma Tun Saroyyan Probolinggo Menggunakan Algoritma Data Mining", Jurnal Teknologi Informasi Indonesia (JTII), 2020

Publication

6%

Exclude quotes On

Exclude bibliography On

Exclude matches < 5%