

IMPLEMENTASI ALGORITMA *DECISION TREE* UNTUK KLASIFIKASI PRODUK LARIS

Asmaul Husnah Nasrullah

¹Fakultas Ilmu Komputer, Universitas Ichsan Gorontalo, Gorontalo, Indonesia

¹asmaulhusnah@unisan.ac.id

ABSTRAK

Algoritma *Decision Tree* C4.5 merupakan algoritma yang dapat digunakan untuk membentuk pohon keputusan. Pohon keputusan (*Decision Tree*) adalah salah satu metode yang cukup mudah untuk diinterpretasikan oleh manusia. Namun algoritma ini belum pernah diuji cobakan untuk klasifikasi produk laris dengan menggunakan data privat (data stok dan penjualan barang di PT Cipta Karya Gorontalo). Oleh karena itu penelitian ini bertujuan untuk menguji akurasi dari algoritma C4.5 dalam melakukan klasifikasi produk laris (data privat). Adapun hasil akurasi model klasifikasi produk laris menggunakan *Decision Tree* C4.5 yang diperoleh dari penelitian ini adalah sebesar 90% dan nilai AUC 0.709 dimana nilai ini termasuk dalam *Good Classification*. Sehingga dapat disimpulkan bahwa model klasifikasi data mining Algoritma *Decision Tree* C4.5 akurat dalam melakukan klasifikasi untuk produk laris.

Kata Kunci— *Decision Tree, C4.5, Klasifikasi, Produk Laris*

ABSTRACT

Decision Tree C4.5 algorithm is an algorithm that can be used to make a decision tree. Decision tree (Decision Tree) is one method that is quite easily interpreted by humans. However, this algorithm has never been tested for product classification using private data (stock data and sales of goods at PT Cipta Karya Gorontalo). Therefore this study aims to test the accuracy of C4.5 in classifying best-selling products (private data). As a result of the evaluation of product classification models using Decision Tree C4.5 obtained from this study amounted to 90% and AUC value of 0.709 where this value is included in the Good Classification. It can be used as a data mining classification method Decision Tree C4.5 algorithm is accurate in classifying hot-selling products.

Keywords— *Decision Tree, C4.5, Classification, Best-Selling Product*

1. PENDAHULUAN

Data Mining merupakan serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual. Sehingga Data Mining menjadi alat yang semakin penting untuk mengubah data tersebut menjadi informasi [1]. Salah satu konferensi internasional terbesar tentang data mining, *IEEE International Conference on Data Mining (ICDM)* menominasikan sepuluh (10) algoritma data mining terbaik yaitu *C4.5*, *K-Means*, *Support Vector Machines (SVM)*, *Apriori*, *Expectation Maximisation Algorithm (EM)*, *Page Rank*, *AdaBoost*, *K-Nearest Neighbors (KNN)*, *Naïve Bayes*, *Classification* dan *Regression Tree (CART)* [2]. Dari ke sepuluh algoritma tersebut tidak disebutkan algoritma yang paling baik. Oleh sebab itu ini menjadi salah satu alasan penelitian dilakukan.

Algoritma *C4.5* merupakan algoritma yang dapat digunakan untuk membentuk pohon keputusan. Pohon keputusan (*Decision Tree*) adalah salah satu metode yang cukup mudah untuk diinterpretasikan oleh manusia. Pohon keputusan adalah model prediksi menggunakan struktur pohon atau struktur berhirarki. Konsep dari pohon keputusan adalah mengubah data menjadi pohon keputusan dan aturan-aturan keputusan [3]. Manfaat utamanya adalah *mem-break down* proses pengambilan keputusan yang kompleks menjadi lebih simpel sehingga pengambil keputusan akan lebih mudah menemukan solusi dari permasalahan yang ada.

Adapun masalah yang dijadikan objek penelitian yaitu produk laris diperusahaan dagang Cipata Karya Gorontalo. Pemilihan produk laris sebagai objek penelitian berdasarkan adanya kesulitan yang dialami oleh perusahaan dagang dalam menentukan produk laris. Kesulitan tersebut disebabkan karena dalam penentuan produk laris perusahaan dagang harus melakukan riset produk. Untuk melakukan riset produk memerlukan waktu dan biaya yang besar, data mining dalam hal ini klasifikasi dapat menjadi solusi untuk mengelompokkan produk laris dan tidak laris tanpa membutuhkan waktu dan biaya yang besar.

Data mining adalah teknik analisis data berbasis pada aplikasi statistik yang bertujuan untuk mengekstrak informasi [4]. Dengan data mining kumpulan data dalam jumlah besar dapat dijadikan informasi lain yang bermanfaat. Data mining dapat melakukan pekerjaan seperti memperkirakan, mengklasifikasikan sampai mengelompokkan data [5]. Menurut Larosse dalam bukunya, dalam mengelompokkan data (klasifikasi) terdapat empat (4) komponen dasar yaitu kelas, prediktor, *training set* dan pengujian *dataset* [6].

Variabel yang digunakan dalam melakukan klasifikasi produk laris yaitu jenis barang, merek dan harga dengan kelas target yaitu laris dan tidak laris. Ketiga variabel ini menjadi tolak ukur yang digunakan dalam riset produk laris di perusahaan dagang. Adapun jenis barang dan merk merupakan variabel bebas sedangkan harga merupakan variabel dependen (terikat). Penelitian ini diharapkan mampu memberikan hasil klasifikasi berupa pohon keputusan (*decision tree*) yang dapat digunakan sebagai bahan pertimbangan perusahaan dagang dalam menentukan pasokan produk berdasarkan jenis, merk dan harga barang yang akan diperdagangkan.

2. TINJAUAN PUSTAKA

2.1. Decision Tree

Pohon (*tree*) adalah sebuah struktur data yang terdiri dari simpul (*node*) dan rusuk (*edge*). Simpul pada sebuah pohon dibedakan menjadi tiga, yaitu simpul akar (*root/node*), simpul percabangan/internal (*branch/internal node*) dan simpul daun (*leaf node*) [3].

Pohon keputusan merupakan representasi sederhana dari teknik klasifikasi untuk sejumlah kelas berhinga, dimana simpul internal maupun simpul akar ditandai dengan nama atribut, rusuk-rusuknya diberi label nilai atribut yang mungkin dan simpul daun ditandai dengan kelas-kelas yang berbeda [3].

2.2. Decision Tree C4.5

Algoritma *C4.5* yaitu sebuah algoritma yang digunakan untuk membentuk *decision tree* (pohon keputusan). Pohon keputusan merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal. Algoritma *C4.5* mengkonstruksi pohon keputusan dari data pelatihan, yang berupa kasus-kasus atau *record-record* (tupel) dalam basis data. Setiap kasus berisikan nilai dari atribut-atribut untuk sebuah kelas. Setiap atribut dapat berisi data diskret atau kontinu (*numerik*). Algoritma *C4.5* juga menangani kasus yang tidak memiliki nilai untuk sebuah atau lebih atribut. Akan tetapi, atribut kelas hanya bertipe diskret dan tidak boleh kosong [8]. Secara umum algoritma *C4.5* untuk membangun pohon keputusan adalah sebagai berikut:

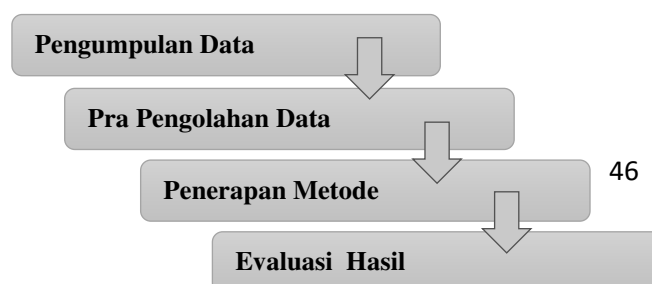
1. Pilih atribut sebagai akar
2. Buat cabang untuk masing-masing nilai
3. Bagi kasus dalam cabang
4. Ulangi proses untuk masing-masing cabang sampai semua kasus pada cabang memiliki kelas yang sama.

2.3. Klasifikasi

Klasifikasi adalah proses dari mencari suatu himpunan model (fungsi) yang dapat mendeskripsikan dan membedakan kelas-kelas data atau konsep-konsep, dengan tujuan dapat menggunakan model tersebut untuk memprediksi kelas dari suatu objek yang mana kelasnya belum diketahui [4].

3. METODE YANG DIUSULKAN

Klasifikasi adalah proses menemukan kumpulan pola atau fungsi-fungsi yang mendeskripsikan dan memisahkan kelas data satu dengan lainnya, untuk dapat digunakan dalam memprediksi data yang belum memiliki kelas data tertentu [7]. Tahapan penelitian untuk model klasifikasi produk laris ditunjukkan pada Gambar 1.



Gambar 1. Tahapan penelitian

1. Pengumpulan Data

Data yang digunakan pada penelitian ini adalah data penjualan dan stok barang yang diperoleh dari UD. Cipta Karya Gorontalo. Data ini merupakan data primer yang diperoleh langsung dari UD. Cipta Karya Gorontalo. Data yang terkumpul sebanyak 450 data dimana data tersebut merupakan data stok dan penjualan barang UD. Cipta Karya Gorontalo yang diambil dari tahun 2015-2018. Data tersebut berisi atribut no transaksi, tgl transaksi, jenis barang, nama barang, merk, harga, jumlah keluar, dan stok barang.

2. Pra Pengolahan Data

Pra pengolahan data yaitu proses membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan-kesalahan pada data. Selain itu pada proses ini juga dilakukan *enrichment* yaitu proses memperkaya data yang sudah ada dengan data lain yang relevan [4].

Untuk mendapatkan data yang berkualitas, dilakukan beberapa teknik pra pengolahan data :

a. Data Validation

Dilakukan untuk mengidentifikasi dan menghapus data yang ganjil (*outlier/noise*) data yang tidak konsisten, dan data yang tidak lengkap (*missing value*)[8]. Data stok dan penjualan barang yang dijadikan dataset akan diidentifikasi apakah ada data yang ganjil, data yang tidak konsisten dan data yang tidak lengkap.

b. Data Cleaning

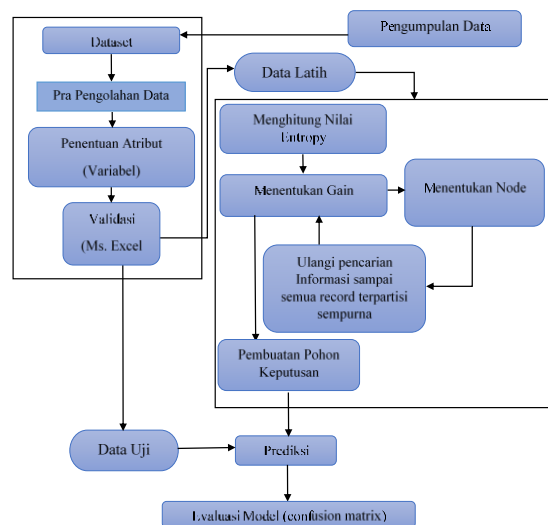
Proses *cleaning* ini mencakup antara lain; membuang duplikasi data, memeriksa data yang inkonsisten dan memperbaiki kesalahan pada data[9]. Dari 450 data yang dikumpulkan akan di analisa apakah ada data yang tidak konsisten atau tidak relevan sehingga akan mengganggu pola aturan dari algoritma yang akan dibentuk.

3. Penerapan Metode

Decision Tree adalah sebuah struktur pohon, dimana setiap simpul (*node*) pohon merepresentasikan atribut yang telah diuji. Setiap cabang merupakan pembagian hasil uji dan *node* daun (*leaf*) merepresentasikan kelompok kelas tertentu. Level *node* teratas dari sebuah *decision tree* adalah akar (*root*) yang biasanya berupa atribut yang paling memiliki pengaruh terbesar pada suatu kelas tertentu. Pada umumnya *Decision Tree* melakukan strategi pencarian secara *top-down* untuk solusinya. Pada proses klasifikasi nilai atribut akan diuji dengan cara mempelajari jalur dari

node akar (*root*) sampai ke *node* akhir (*leaf*) baru kemudian kelas baru akan ditentukan [10].

Model klasifikasi produk laris yang diusulkan pada penelitian ini dapat dilihat pada Gambar 2.



Gambar 2. Model yang diusulkan

4. Evaluasi Hasil

Setelah hasil klasifikasi diperoleh, selanjutnya hasil di evaluasi menggunakan *Cross Validation (confusion matrix)* untuk melihat akurasi, presisi dan *recall* yang dihasilkan oleh model yang diusulkan.

4. HASIL PENELITIAN

Algoritma C4.5 merupakan salah satu algoritma yang digunakan untuk melakukan klasifikasi atau segmentasi yang bersifat prediktif.[11]. Algoritma C4.5 digunakan untuk membentuk pohon keputusan. Pohon keputusan merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal handal. Pohon keputusan mengubah fakta yang sangat besar menjadi aturan-aturan, dengan begitu aturan tersebut dapat dengan mudah dipahami [12].

Menurut kusrini dalam bukunya, secara umum algoritma C4.5 membangun pohon keputusan terdiri dari beberapa tahapan. Berikut garis besarnya [13];

- Pilih atribut sebagai akar
- Buat cabang untuk tiap-tiap nilai
- Bagi kasus dalam cabang
- Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

1. Memilih atribut sebagai simpul akar

Pemilihan atribut yang akan dijadikan akar adalah dengan menghitung nilai *gain* dari semua atribut. Dan

yang dipilih mejadi akar pertama adalah yang memiliki nilai *gain* tertinggi. Namun sebelum menentukan nilai *gain*, terlebih dahulu hitung nilai entropy. Penentuan nilai *entropy* menggunakan persamaan berikut [5];

$$Entropy (S) = \sum_{j=1}^n p_i \cdot \log_2 p_i \quad (1)$$

Keterangan:

S = himpunan kasus

n = jumlah partisi

pi = proporsi Si terhadap S

Setelah itu tentukan nilai gain menggunakan persamaan (2)[11];

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

Keterangan:

S = himpunan kasus

A = atribut

n = jumlah partisi atribut A

|Si| = jumlah kasus pada partisi ke – i

|S| = jumlah kasus dalam s

Pada penelitian ini terdapat empat (4) atribut yaitu jenis barang, merk, harga dan kelas barang. Tabel 1 menunjukkan contoh data training yang berisi atribut dan nilai untuk masing-masing atribut. Data ini selanjutnya digunakan untuk menguji model klasifikasi yang dibuat.

Tabel 1. Data Uji

No Data	Jenis Barang	Merk	Harga	Kelas
1	kramik	Vicenza	385500	Laris
2	kramik	Vicenza	538500	Laris
3	kramik	Vicenza	538500	Laris
4	plastik	Sunlife	18300	Laris
5	plastik	Sunlife	12500	Laris
6	plastik	Fori	27500	Laris
7	kaca	Vicenza	520000	Laris
8	Kaca	Sunlife	179000	Tidak Laris
9	kramik	Sunlife	178000	Laris
10	kramik	Vicenza	44500	Laris
11	Kaca	Vicenza	600000	Laris
12	Kaca	Sunlife	285000	Tidak Laris
13	Kaca	Sunlife	329000	Tidak Laris
14	Kaca	Vicenza	415000	Laris
15	kramik	Vicenza	309000	Laris
16	kramik	Vicenza	865000	Laris
17	kramik	Fori	999000	Laris
18	plastik	Fori	339000	Laris
19	plastik	Fiori	289000	Laris
20	plastik	Fiori	559500	Laris

...				
dst				

Di bawah ini, contoh perhitungan nilai *entropy* dan *gain* untuk setiap atribut :

Entropy Jenis Barang:

$$Entropy (Total, Plastik) = (- \frac{16}{16} * \log_2(\frac{16}{16})) + (- \frac{0}{16} * \log_2(\frac{0}{16})) = 0$$

$$Entropy (Total, Kramik) = (- \frac{17}{17} * \log_2(\frac{17}{17})) + (- \frac{0}{17} * \log_2(\frac{0}{17})) = 0$$

$$Entropy (Total, Kaca) = (- \frac{6}{17} * \log_2(\frac{6}{17})) + (- \frac{11}{17} * \log_2(\frac{11}{17})) = 0,936667$$

Entropy Merek

$$Entropy (Total, Vicenza) = (- \frac{18}{18} * \log_2(\frac{18}{18})) + (- \frac{0}{19} * \log_2(\frac{0}{19})) = 0$$

$$Entropy (Total, Sunlife) = (- \frac{11}{18} * \log_2(\frac{11}{18})) + (- \frac{7}{18} * \log_2(\frac{7}{18})) = 0,964078$$

$$Entropy (Total, Fiori) = (- \frac{12}{14} * \log_2(\frac{12}{14})) + (- \frac{2}{14} * \log_2(\frac{2}{14})) = 0,591673$$

Entropy Harga

$$Entropy (Total, Harga A) = (- \frac{29}{35} * \log_2(\frac{29}{35})) + (- \frac{6}{35} * \log_2(\frac{6}{35})) = 0,660962$$

$$Entropy (Total, Harga B) = (- \frac{10}{15} * \log_2(\frac{10}{15})) + (- \frac{5}{15} * \log_2(\frac{5}{15})) = 0,918296$$

$$Entropy (Total) = (- \frac{39}{50} * \log_2(\frac{39}{50})) + (- \frac{11}{50} * \log_2(\frac{11}{50})) = 0,760167$$

$$Entropy (Total) = 0,760167$$

Selanjutnya diperoleh nilai *gain* untuk masing-masing atribut;

$$Gain (Total, Jenis) = 0,760167 - (\frac{16}{50} * 0) + (\frac{17}{50} * 0) + (\frac{17}{50} * 0,936667) = \mathbf{0,441700}$$

$$Gain (Total, Merek) = 0,760167 - (\frac{18}{50} * 0) + (\frac{18}{50} * 0,964078) + (\frac{14}{50} * 0,591673) = 0,247431$$

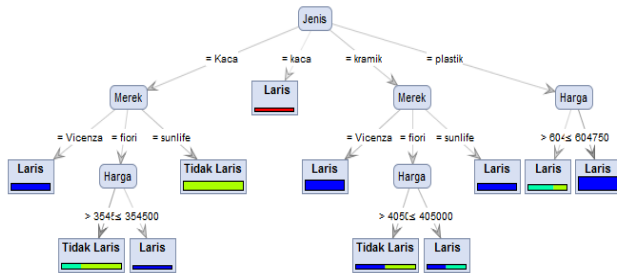
$$Gain (Total, Harga) = 0,760167 - (\frac{35}{50} * 0,660962) + (\frac{15}{50} * 0,918296) = 0,022005$$

Berdasarkan perhitungan diatas, dapat dilihat nilai *gain* tertinggi ada pada atribut jenis barang yaitu 0.4417 maka atribut jenis barang di jadikan akar (*root*) dari pohon keputusan.

2. Membuat cabang untuk setiap nilai

Untuk menentukan simpul selanjutnya, dilakukan perhitungan nilai *entropy* dan *gain* dengan cara yang sama

sehingga diperoleh pohon keputusan seperti gambar di bawah ini:



Gambar 3. Pohon Keputusan hasil Algoritma *Decision Tree C4.5*

Berdasarkan pohon keputusan (*decision tree*) untuk model klasifikasi produk laris diatas. Dapat dilihat bahwa atribut yang mempunyai pengaruh utama dalam penentuan produk laris adalah atribut jenis barang, kemudian merk dan terakhir harga.

Dari pohon keputusan pada gambar 3 didapatkan *rule* untuk klasifikasi produk laris. *Rule* yang didapat sebagai berikut :

```

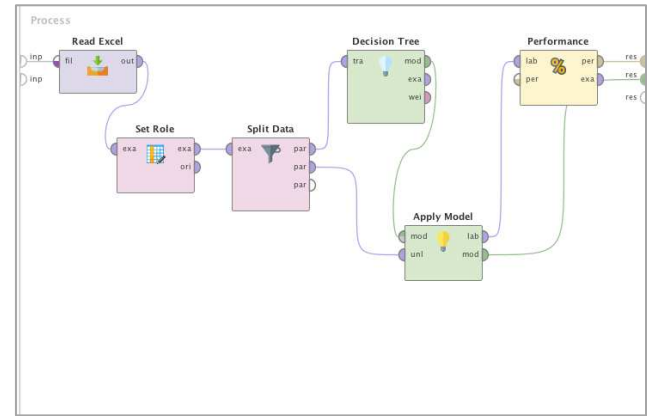
Tree
Jenis = Kaca
| Merek = Vicenza: Laris (Laris=5, Laris=0, Tidak Laris=0, Laris=0)
| Merek = fiori
| | Harga > 354500: Tidak Laris (Laris=0, Laris=1, Tidak Laris=2, Laris=0)
| | Harga ≤ 354500: Laris (Laris=1, Laris=0, Tidak Laris=0, Laris=0)
| Merek = sunlife: Tidak Laris (Laris=0, Laris=0, Tidak Laris=7, Laris=0)
Jenis = kaca: Laris (Laris=0, Laris=0, Tidak Laris=0, Laris=1)
Jenis = kramik
| Merek = Vicenza: Laris (Laris=8, Laris=0, Tidak Laris=0, Laris=0)
| Merek = fiori
| | Harga > 405000: Tidak Laris (Laris=1, Laris=0, Tidak Laris=1, Laris=0)
| | Harga ≤ 405000: Laris (Laris=1, Laris=1, Tidak Laris=0, Laris=0)
| Merek = sunlife: Laris (Laris=5, Laris=0, Tidak Laris=0, Laris=0)
Jenis = plastik
| Harga > 604750: Laris (Laris=0, Laris=2, Tidak Laris=1, Laris=0)
| Harga ≤ 604750: Laris (Laris=13, Laris=0, Tidak Laris=0, Laris=0)
    
```

Gambar 4. *Rule* hasil algoritma *Decision Tree C4.5*

3. Evaluasi Model

Pada penelitian ini setelah hasil klasifikasi dan *rule* berupa pohon keputusan dibuat, selanjutnya dilakukan uji kinerja model klasifikasi dengan *Decision Tree C4.5* dilakukan dengan menggunakan *cross validation* (*Confusion Matrix*) dan *Area Under Cover* (AUC). *Confusion Matrix* memberikan rincian dari kesalahan klasifikasi.

Pengujian ini dilakukan dengan memanfaatkan *tools* RapidMiner, hasil pengujian dapat dilihat pada Gambar 5 dan Gambar 6.



Gambar 5. Model klasifikasi menggunakan algroritma *Decision Tree C4.5*

Tabel 3. *Confusion Matrix*

CLASSIFICATION	PREDICTED CLASS
OBSERVED CLASS	Class = YES
	Class = NO
Class=YES	true positive-TP
	false negatif-FN
Class=NO	false positive-FP
	true negatif-TN

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (3) [14]$$

Berdasarkan persamaan (3) maka diperoleh nilai akurasi untuk model klasifikasi pada Gambar 5 sebagai berikut :

Diketahui:

$$TP = 45$$

$$TN = 0$$

$$\text{Akurasi} = \frac{45+0}{45+0+2+3} = 0.9$$

accuracy: 90.00%			
	true Laris	true Tidak Laris	class precision
pred. Laris	45	3	93.75%
pred. Tidak Laris	2	0	0.00%
class recall	95.74%	0.00%	

Gambar 6. Hasil evaluasi menggunakan *cross validation*

Gambar 6 menunjukkan nilai akurasi yang sama dengan nilai akurasi yang dihasilkan dari perhitungan manual diatas yaitu sebesar 0.90 atau 90%.

Receiver Operating Characteristic (ROC) Curve adalah grafik antara sensitifitas (*true positive rate*) pada sumbu Y dengan 1 spesifitas pada sumbu X (*false positive rate*), yang tujuannya untuk menentukan *cut off point* pada uji diagnostik yang bersifat kontinyu[15]. Untuk klasifikasi data mining, nilai AUC dapat dibagi menjadi beberapa kelompok [16];

1. AUC bernilai 0.90 – 1.00 = *Excellent classification*
2. AUC bernilai 0.80 – 0.90 = *Good classification*
3. AUC bernilai 0.70 – 0.80 = *Fair classification*
4. AUC bernilai 0.60 – 0.70 = *Poor classification*
5. AUC bernilai 0.50 – 0.60 = *Failure*

5. KESIMPULAN

Dari hasil penelitian yang telah dilakukan untuk data penjualan dan stok barang di UD. Cipta Karya Gorontalo maka diperoleh akurasi sebesar 90% dan nilai AUC 0.709. Dimana nilai ini termasuk dalam *Fair Classification* (Klasifikasi yang cukup). Sehingga dapat disimpulkan bahwa model klasifikasi data mining menggunakan Algoritma *Decision Tree C4.5* akurat dalam melakukan klasifikasi untuk produk laris.

Namun jika ditelusuri lebih lanjut ternyata model klasifikasi ini masih belum bisa dinyatakan sebagai model klasifikasi terbaik untuk klasifikasi produk laris dengan data yang diperoleh dari UD Cipta Karya Gorontalo. Oleh karena itu berikut beberapa saran yang bisa dijadikan sebagai acuan untuk penelitian lebih lanjut;

1. Diperlukan teknik pra pengolahan data dan transformasi data yang lebih baik dengan cara memperbaiki inkonsistensi data pada *dataset*.
2. Penelitian ini dapat dikembangkan dengan mengoptimasi model klasifikasi yang telah dibuat dengan menggunakan *Particle Swam Optimization (PSO)* atau *Genetic Algorithm (GA)*

UCAPAN TERIMA KASIH

Ucapan terima kasih kami haturkan kepada Usaha Dagang Cipta Karya Gorontalo yang telah bersedia memberikan dan

menyediakan data stok dan penjualan barang, yang kemudian dijadikan *dataset* pada penelitian ini.

Daftar Pustaka

- [1] B. Utami and P. Aliandu, "KLASIFIKASI PENENTUAN TIM UTAMA OLAHRAGA HOCKEY MENGGUNAKAN ALGORITMA C4.pdf," *Proc. Int. Conf. Information, Commun. Technol. Syst.*, vol. 5, no. 4, pp. 1–5, 2013.
- [2] X. Wu *et al.*, *Top 10 algorithms in data mining*, vol. 14, no. 1. 2008.
- [3] J. Eska, "Data Mining Untuk Prediksi Penjualan Wallpaper Menggunakan Algoritma C45," *JURTEKSI (Jurnal Teknol. dan Sist. Informasi)*, vol. 2, pp. 9–13, 2016.
- [4] D. F. Ristianti, "Komparasi Algoritma Klasifikasi pada Data Mining," vol. 1, no. 1, pp. 148–156, 2019.
- [5] Y. Mardi, "Data Mining : Klasifikasi Menggunakan Algoritma C4.5," *J. Edik Inform.*, vol. 2, no. 2, pp. 213–219, 2017.
- [6] W. D. Septiani, "Dan Naive Bayes Untuk Prediksi Penyakit Hepatitis," vol. 13, no. 1, pp. 76–84, 2017.
- [7] A. H. Nasrullah, "Penerapan Metode C4.5 untuk Klasifikasi Mahasiswa Berpotensi Drop Out," *Ilk. J. Ilm.*, vol. 10, no. 2, pp. 244–250, 2018, doi: 10.33096/ilkom.v10i2.300.244-250.
- [8] I. Carolina and R. Kresna, "Klasifikasi kelahiran prematur menggunakan algoritma c4.5," *Semin. Nas. Teknol.*, pp. 668–672, 2018.
- [9] Hariati, M. Wati, and B. Cahyono, "Penerapan Algoritma C4.5 Decision Tree pada Penentuan Penerima Program Bantuan Pemerintah Daerah Kabupaten Kutai Kartanegara," *Jurti*, vol. 2, no. 1, pp. 27–36, 2018.
- [10] Y. Rosela, "IMPLEMENTASI KLASIFIKASI DECISION TREE MENGANALISA STATUS PENJUALAN BARANG MENGGUNAKAN C4.5 (Studi Kasus : Pt . Matahari Department Store Medan Mall)," *J. Pelita Inform.*, vol. 18, no. 1, pp. 143–150, 2019.
- [11] D. R. S. P, A. P. Windarto, D. Hartama, and I. S. Damanik, "Penerapan Klasifikasi C4.5 Dalam Meningkatkan Sistem Pembelajaran Mahasiswa," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 3, no. 1, pp. 593–597, 2019, doi: 10.30865/komik.v3i1.1665.
- [12] D. R. Ente, S. A. Thamrin, S. Arifin, H. Kuswanto, and A. Andreza, "Klasifikasi Faktor-Faktor Penyebab Penyakit Diabetes Melitus Di Rumah Sakit Unhas Menggunakan Algoritma C4.5," *Indones. J. Stat. Its Appl.*, vol. 4, no. 1, pp. 80–88, 2020, doi: 10.29244/ijsa.v4i1.330.

- [13] L. N. Rani, "Klasifikasi Nasabah Menggunakan Algoritma C4.5 Sebagai Dasar Pemberian Kredit," *INOVTEK Polbeng - Seri Inform.*, vol. 1, no. 2, p. 126, 2016, doi: 10.35314/isi.v1i2.131.
- [14] D. Istiawan and L. Khikmah, "Implementation of C4.5 Algorithm for Critical Land Prediction in Agricultural Cultivation Areas in Pemali Jratun Watershed," *Indones. J. Artif. Intell. Data Min.*, vol. 2, no. 2, p. 67, 2019, doi: 10.24014/ijaidm.v2i2.7569.
- [15] P. Assiroj, "Kajian Perbandingan Teknik Klasifikasi Algoritma C4 . 5 , Naïve Bayes Dan Cart Untuk Prediksi Kelulusan Mahasiswa (Studi Kasus : Stmik Rosma Karawang) (Studi Kasus : Stmik Rosma Karawang)," *Media Inform.*, vol. 15, no. 2, pp. 1–17, 2018, doi: 10.5281/zenodo.1184054.
- [16] E. Fitriani, R. Aryanti, A. Saepudin, and D. Ardiansyah, "Penerapan Algoritma C4.5 Untuk Klasifikasi Penempatan Tenaga Marketing," *Paradig. - J. Komput. dan Inform.*, vol. 22, no. 1, pp. 72–78, 2020, doi: 10.31294/p.v22i1.6898.