

PENERAPAN ALGORITMA C4.5 UNTUK PREDIKSI AKURASI KEPUASAN PELANGGAN PO. SINAR JAYA

C4.5 ALGORITHM APPLICATION FOR PREDICTION OF CUSTOMER SATISFACTION ACCURACY IN PO. SINAR JAYA

Ismasari Nawangsih¹, Amirudin²

Program Studi Teknik Informatika, Sekolah Tinggi Teknologi Pelita Bangsa.
Jl. Inspeksi Kalimalang Jl. Tegal Danas. Cikarang Pusat, Bekasi, Jawa Barat 17530
Telp : (021) 28518181
ismasari.n@pelitabangsa.ac.id

ABSTRAKS

Kepuasan pelanggan sangat dibutuhkan untuk berlangsungnya suatu bidang usaha, salah satunya bidang usaha jasa transportasi. Apalagi sekarang ini begitu banyak persaingan yang terjadi, perlu adanya pengetahuan untuk menentukan tingkat kepuasan pelanggan sebagai keputusan untuk memberikan pelayanan yang maksimal. Karena itu dibutuhkan teknik klasifikasi dan prediksi data mining dalam memecahkan masalah tersebut. Klasifikasi yang digunakan dalam data mining adalah Decision Tree dikarenakan merupakan teknik yang banyak digunakan dan menghasilkan output yang dengan aturan yang ada, dengan hal tersebut dapat menyajikan data kepuasan pelanggan. Dalam penelitian ini dilakukan klasifikasi dengan metode algoritma c4,5 dan menghasilkan aturan klasifikasi kepuasan pelanggan dengan akurasi 85%.

Kata kunci : akurasi kepuasan pelanggan, algoritma c4.5, decision tree, data mining.

ABSTRACT

Customer satisfaction is needed for the ongoing business field, one of them is the transportation service sector. Especially now that there is so much competition, there needs to be knowledge to determine the level of customer satisfaction as a decision to provide maximum service. because it requires data mining classification and prediction techniques in solving these problems. Classification used in data mining is a Decision Tree because it is a technique that is widely used and produces output with existing rules, with it can present customer satisfaction data. In this study classification was carried out using the c4.5 algorithm method and resulted in the classification rules of customer satisfaction with 85% accuracy.

Keywords: accuracy of customer satisfaction, c4.5 algorithm, decision tree, data mining

1. PENDAHULUAN

Bisnis jasa transportasi semakin meningkat dan persaingan bersaing dengan cepat. Maka Untuk membantu meningkatkan pelanggan transportasi di P.O Sinar jaya . Penulis ingin melakukan pengolahan informasi data kepuasan konsumen, yang diperoleh penulis dengan melakukan penyebaran kuisioner kepada konsumen bus Sinar Jaya. Metode yang diterapkan penulis dalam mengolah informasi data tersebut, ialah menggunakan metode algoritma C4.5.

Secara umum permasalahan dalam penelitian ini adalah untuk mengetahui tingkat kepuasan konsumen tentang pelayanan dan fasilitas bus Sinar Jaya, dan belum adanya suatu metode akurasi dalam memprediksi Kepuasan pelanggan.

Berdasarkan permasalahan yg ada maka merumuskan masalah menjadi : Bagaimana penerapan algoritma c4.5 untuk prediksi akurasi

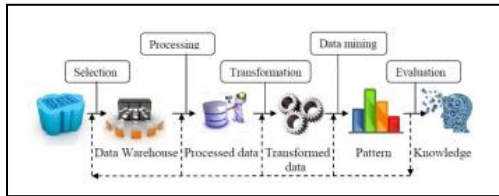
kepuasan pelanggan PO. Sinar Jaya ?; dan Bagaimana mengetahui tingkat kepuasan pelanggan PO. Sinar Jaya?

Data Mining

Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *database* besar. (Santoso, 2007)

Tahap – Tahap Data Mining

Sebagai suatu rangkaian proses, data mining dapat dibagi menjadi beberapa tahap. Tahap-tahap tersebut bersifat interaktif di mana pemakai terlibat langsung atau dengan perantaraan knowledge base. (Kursini, Emha, 2009)



Gambar 1. Tahap Data Mining
 Sumber : Saputra Ilham.(2016)-19

Metode Data Mining

Data Mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu : Deskripsi, Estimasi, Prediksi, Klasifikasi, Pengklusteran, dan Asosiasi.

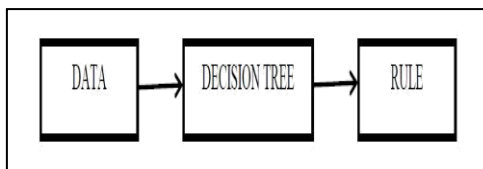
Klasifikasi merupakan proses menemukan sebuah model atau fungsi yang mendeskripsikan dan membedakan data ke dalam kelas-kelas. Klasifikasi melibatkan proses pemeriksaan karakteristik dari objek dan memasukkan objek ke dalam salah satu kelas yang sudah didefinisikan sebelumnya. (Larose, 2005) proses klasifikasi didasarkan pada empat komponen mendasar, yaitu :

- a. Kelas(Class)
- b. Prediktor(Predictor)
- c. Pelatihan dataset(Training dataset)
- d. Dataset pengujian(Testing dataset)

Decision Tree Algoritma C4.5

Pohon keputusan adalah salah satu metode klasifikasi yang kuat dan terkenal. Metode *Decision Tree* mengubah fakta besar menjadi pohon keputusan yang mewakili aturan, sehingga aturan tersebut dapat dengan mudah dipahami oleh manusia. *Decision Tree* juga berfungsi untuk mengeksplorasi data, menemukan hubungan tersembunyi antara sejumlah variabel input dan variabel tujuan. (Undavia J N, Patel dan Donia, 2013)

Konsep dari *Decision tree* adalah mengubah data menjadi pohon keputusan dan aturan keputusan.



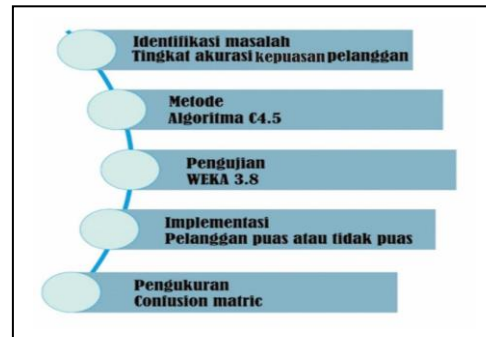
Gambar2 Konsep Decision tree

Weka

WEKA (Waikato Environment for Knowledge Analysis) merupakan Sebuah perangkat lunak open source yang menyediakan berbagai metode data mining yang siap digunakan . Weka berbasis framework Java sehingga dapat digunakan di berbagai macam operasi.(Adinugroho, Sigit dan Sari, Yuita Arum, 2018)

2. METODE PENELITIAN

Metode penelitian penulis di Po. Sinar Jaya yang bergerak dalam Jasa Transportasi . Penulis meneliti masalah dan melakukan penyelesaian dengan menggunakan metode *data mining decision tree* algoritma c4.5. Kemudian melakukan pengujian data pada *tools* WEKA, mengimplementasikan terhadap tingkat akurasi kepuasan pelanggan dengan menghitung hasil *confusion matrix*.



Gambar3.Langkah Metode Penelitian PO.Sinar Jaya

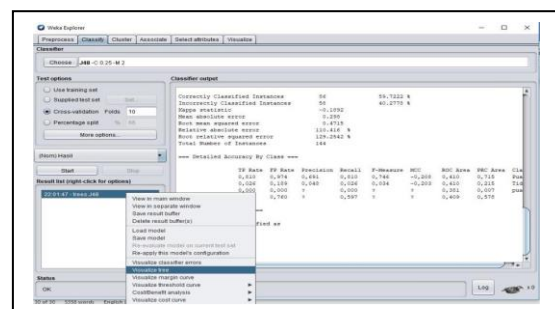
3. HASIL DAN PEMBAHASAN

Proses Data Mining

Pada pembahasan sebelumnya telah dijelaskan secara teori bagaimana proses *data mining* klasifikasi algoritma C4.5, dijelaskan juga bagaimana perhitungan rumus algoritma C4.5 untuk mencari sebuah pohon keputusan dari data yang di proses. Dan kali ini akan dijelaskan proses penerapan *data mining* pada WEKA.

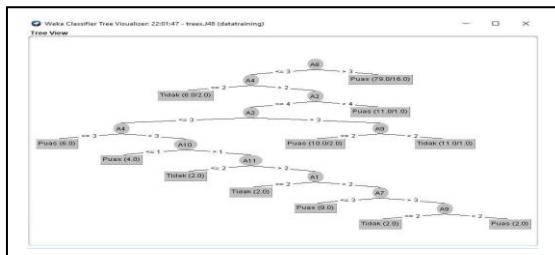
Proses Data Training Pada WEKA

Data training disini adalah data dengan presentase 80 % dari data keseluruhan 144 data. Data training yang akan diproses pada WEKA, ialah data training .



Gambar 4 Hasil Proses Mining

Dan berikut *visual tree* yang dihasilkan berikut dengan penjelasan *rule*-nya.



Gambar 5 Hasil Pohon Keputusan

Dari proses pada WEKA data training yang diproses menghasilkan 21 data error, dengan presentase 80% data training memiliki data sebanyak 144 data. Hal ini diketahui dari hasil *confusion matrix*, berikut tabel pengetahuan dari hasil *confusion matrix*.

Tabel 1. Hasil Testing Data Training

Jumlah Data Training	Jumlah Error	Positif benar (TP)	Positif salah (FP)	Negatif salah (FN)	Negatif benar (TN)
144	21	103	3	18	20

Penjelasan tabel diatas ialah dari 144 data yang diuji terdapat kasus error 21 data, yaitu kasus dengan hasil dan prediksi yang tidak sama. TP (*true positive*) ialah data dengan hasil puas dan prediksinya juga puas. FP (*false positive*) ialah data dengan hasil puas namun prediksinya tidak puas. FN (*false negative*) ialah data dengan hasil tidak puas namun prediksinya puas. Dan TN (*true negative*) ialah data dengan hasil tidak puas dan prediksi tidak puas.

Tabel 2. Hasil Pengujian Data Training.

Jenis Data	Presentase	Jumlah Data	Akurasi	Precision	Recall
Data Training	80 %	144	85 %	97 %	85 %

Data Testing dan Pengujian Data Testing

Data testing merupakan data yang nantinya untuk menguji atau mengevaluasi aturan yang dihasilkan dari pohon keputusan. Data training memiliki presentase 20 % dari keseluruhan data yaitu sebanyak 36 data, dengan hasil puas sebanyak 28 data dan 8 data dengan hasil tidak puas.

Tabel 3. Hasil Testing Data Testing

Jumlah data testing	Jumlah Error	Positif benar (TP)	Positif salah (FP)	Negatif salah (FN)	Negatif benar (TN)
36	19	15	13	6	2

Dengan melakukan pengujian dan menghitung akurasi, precision dan recall. Didapatkan sebuah tabel hasil sebagai berikut :

Tabel 4. Evaluasi dan Validasi

Jenis Data	Presentase	Jumlah Data	Akurasi	Precision	Recall
Data Testing	20 %	36	42 %	83 %	71 %

Berdasarkan tabel diatas diketahui hasil akurasi dari data training adalah 85 %, dengan nilai precision sebesar 97 % dan recall sebesar 85 %. Untuk data testing diketahui nilai akurasi sebesar 42 %, nilai precision sebesar 83 %, dan recall sebesar 71 %.

Hasil Tingkat Kepuasan Pelanggan

Dari data sampel yang diutarakan yaitu sebanyak 180 data responden, terdapat 134 data responden yang menyatakan puas terhadap layanan Po. Sinar Jaya. Dan ada 46 data responden yang menyatakan tidak puas, dapat ditentukan tingkat kepuasan dari pengetahuan tersebut.

Tabel 5. Tingkat Kepuasan Pelanggan

No	Respon	Tingkat Presentase
1	Konsumen puas	74.5 %
2	Konsumen tidak puas	25.5%

Dihasilkan sebuah tingkat kepuasan pelanggan dengan data sampel yang ada yaitu sebesar 74,5 %. Dan tingkat ketidakpuasan pelanggan sebesar 25,5% atau bisa dikatakan dua pertiga dari data sampel menyatakan kepuasannya terhadap pelayanan PO. Sinar Jaya.

4. KESIMPULAN

1. Evaluasi hasil klasifikasi dalam prediksi kompetensi menggunakan data training dengan *confusion matrix* tingkat akurasi sebesar 85 % dan data testing tingkat akurasi sebesar 42 %.
2. Dengan mengimplementasikan data responden sebanyak 180 data, didapatkan tingkat kepuasan pelanggan PO. Sinar Jaya sebesar 74.5%.

DAFTAR PUSTAKA

Adinugroho, Sigit., dan Sari, Yunita Arum. 2018. Impelementasi Data Mining Menggunakan Weka. Malang : UB Press.

Agustina Vina. 2012. Analisis Pengaruh Kualitas Pelayanan, Kepuasan Pelanggan, Dan Nilai Pelanggan Dalam Meningkatkan Loyalitas Pelanggan Joglosemar Bus (Studi Pada Wilayah Semarang *Town Office*),1-64.

Chapman P. 2000.CRISP-DM 1.0: Step-by-step Data Mining Guide, SPSS.

Gorunescu, F. 2011. Data mining: Concepts, models

- and techniques. *Intelligent Systems Reference Library*, 12. <http://doi.org/10.1007/978-3-642-19721-5>
- Harry Dhika, Fitriana Destiawati dan Aswin Fitrians. 2016. Implementasi Algoritma C4.5 Terhadap Kepuasan Pelanggan, 16–22.
<https://www.cs.waikato.ac.nz/ml/weka/downloadi ng.html> (diakses tanggal 07 Oktober 2018)
<https://www.sinarjayagroup.co.id/> (diakses 18 Oktober 2018)
- Kamagi, D. H., dan Hansun, S. 2014. Implementasi Data Mining dengan Algoritma C4 . 5 untuk Memprediksi Tingkat Kelulusan Mahasiswa. *ultimatics, Vol. VI, No. 1 | Juni 2014, VI(1)*, 15–20. <https://doi.org/10.31937/ti.v6i1.327>.
- Kusrini. 2009. *Algoritma Data Mining*. Yogyakarta: Andi.
- Lakshmi, T. M., Martin, A., Begum, R. M., dan Venkatesan, V. P. 2013. An Analysis on Performance of Decision Tree Algorithms using Student’s Qualitative Data. *International Journal of Modern Education and Computer Science*, 5(5), 18–27.
<https://doi.org/10.5815/ijmeecs.2013.05.03>
- Larose, D. T. 2005. *Discovering knowledge in data: An introduction to data mining*. New Jersey: John Willey & Sons, Inc.
- Martiningsih, F., Informatika, P. S., Komunikasi, F., Informatika, D. A. N., & Surakarta, U. M. 2018. Sistem evaluasi kepuasan pelanggan go-jek menggunakan metode naïve bayes.
- Rambat Lupiyoadi. 2004. *Manajemen Pemasaran Jasa : Teori dan Pratek*, Jakarta : PT Salemba Empat.
- Rohman, F. I. 2016. Penerapan Algoritma C4 . 5 Pada Kepuasan Pelanggan Perum DAMRI. *Ibnu Ftchur Rohman*, 2(1), 1–14.
- Santoso, T. B. 2015. Analisa Dan Penerapan Metode C4.5 Untuk Prediksi Loyalitas Pelanggan. *CEUR Workshop Proceedings*, 1542(1), 33–36. <https://doi.org/10.1017/CBO9781107415324.004>
- Subiyanto, S., Negeri, U., & View, E. E. 2017. Recommendation System for Vocational Major Streaming By C4 . 5, (June).
- Sunge, A. S., Studi, P., Informatika, T., Tinggi, S., Pelita, T., Pusat, C., ... Masalah, R. 2018. Prediksi Kompetensi Karyawan Menggunakan Algoritma C4 . 5 (Studi Kasus : Pt Hankook Tire Indonesia), 2018(Sentika), 23–24.
- Syukria Darman. 2012. Pengaruh Kualitas Pelayanan Jasa Transportasi Taksi Terhadap Loyalitas Konsumen Pada PT Putra Transpor Nusantara Bandung. Bandung, Universitas Widyatama.
- Tjjiptono Fandy.1997.Prinsip – Prinsip Total Quality Service Yogyakarta: Andi Offset.
- Undavia J N, Patel dan Dolia. 2013. “Comparison of Classification Algorithms to Predict Students’ Post Graduation Course in Weka Environment. *International Journal of Advanced Research in Computer Science and Software Engineering*. vol. 3. no. 9.