

Penerapan Data Mining Untuk Klasifikasi Kepuasan Pelanggan Transportasi Online (Ojek Online) Menggunakan Algoritma C.4.5

Application of Data Mining for Online Transportation Customer Satisfaction

Classification (Ojek Online) Using the C.4.5 Algoritma Algorithm

Suherman¹, Donny Maulana², Vivi Mustikaningtyas³

^{1,2,3}Program Studi Teknik Informatika, Fakultas Teknik, Universitas Pelita Bangsa

¹suherman@pelitabangsa.ac.id, ²donny.maulana@pelitabangsa.ac.id, ³vivirozikin29@gmail.com*

Abstract

Human activities to meet their daily needs and needs, both at work, or just for a walk. So, this needs to be supported by adequate transportation. With the development of technology today there are applications that introduce motorcycle taxi booking services using technology and use service standards. In Indonesia there are many motorbikes, which also function as general vehicles, namely transporting people / goods. Currently there are many online transportation service providers (online motorcycle taxi) known as Go-Jek, Grab, and Uber. Customer satisfaction input attributes in this study include price, facilities, service and loyalty. Data mining is a series of processes to explore added value in the form of information that has not been known manually from a database. In this study it is expected to help the online transportation services in increasing customer satisfaction. Based on the results of the classification using C4.5 algorithm shows that the accuracy reached 75.33%, which shows that the C4.5 algorithm is suitable for measuring the level of satisfaction of online transportation.

Keywords: *Satisfaction of online transportation, Data Mining, Algoritma C.4.5*

Abstrak

Aktivitas manusia untuk memenuhi kebutuhan dan kebutuhannya sehari-hari, baik di tempat kerja, maupun sekedar jalan-jalan. Sehingga perlu didukung dengan transportasi yang memadai. Dengan berkembangnya teknologi saat ini terdapat aplikasi yang memperkenalkan layanan pemesanan ojek dengan menggunakan teknologi dan menggunakan standar pelayanan. Di Indonesia banyak terdapat sepeda motor yang juga berfungsi sebagai kendaraan umum yaitu mengangkut orang/barang. Saat ini sudah banyak penyedia jasa transportasi online (ojek online) yang dikenal dengan nama Go-Jek, Grab, dan Uber. Atribut input kepuasan pelanggan dalam penelitian ini meliputi harga, fasilitas, pelayanan dan loyalitas. Data mining merupakan rangkaian proses untuk menggali nilai tambah berupa informasi yang belum diketahui secara manual dari suatu database. Dalam penelitian ini diharapkan dapat membantu layanan transportasi online dalam meningkatkan kepuasan pelanggan. Berdasarkan hasil klasifikasi menggunakan algoritma C.4.5 menunjukkan akurasi mencapai 75,33% yang menunjukkan bahwa algoritma C.4.5 cocok untuk mengukur tingkat kepuasan transportasi online.

Kata kunci: Kepuasan Pelanggan Transportasi Online, Data Mining, Algoritma C4.5

Pendahuluan

Aktivitas manusia untuk memenuhi keperluan dan kebutuhan hidupnya sehari-hari, baik dalam bekerja, ataupun sekedar jalan-jalan. Maka, hal ini perlu didukung dengan transportasi yang memadai. Dengan perkembangannya teknologi saat ini terdapat aplikasi yang mengenalkan layanan pemesanan ojek menggunakan teknologi dan memakai standar pelayanan. Semakin hari manusia semakin kreatif dan inovatif menciptakan bisnis baru, terutama bisnis yang menyediakan jasa transportasi. karena manusia di zaman sekarang membutuhkan transportasi yang cepat, nyaman, aman dan [1] penduduknya yang lumayan padat, kemacetan sering terjadi dan masyarakat banyak yang menggunakan kendaraan pribadi sehingga muncul kemacetan tersebut, sehingga transportasi umum lebih jenuh ketika menghadapi kemacetan yang terjadi.

Pengusaha jasa transportasi berlomba-lomba memikat konsumennya dengan peningkatan pelayanan, kemudahan pemesanan, kenyamanan armada, ketepatan waktu dan lain sebagainya. Di Indonesia banyak di jumpai sepeda motor, yang juga berfungsi sebagai kendaraan umum yaitu mengangkut orang/barang. Saat ini sudah banyak penyedia jasa transportasi online (ojek online) yang dikenal dengan nama Go-Jek, Grab, dan Uber. Semua memberikan pelayanan yang hampir sama mulai dari mengantarkan orang dengan biaya yang berbeda-beda, namun dengan sistem pemesanan yang sama yaitu pemesanan melalui aplikasi telephone genggam maupun website.

Menurut Muhamad Rizki Ilham bahwa Kualitas jasa dan Pelayanan merupakan salah satu unsur yang sangat penting dalam menciptakan keputusan konsumen. Salah satu cara untuk menempatkan hasil pelayanan yang lebih unggul daripada pesaing adalah dengan memberikan pelayanan yang baik, efisien, dan cepat [1]. Menurut Panjaitan, Yulianti bahwa Kualitas jasa pada umumnya tidak dilihat dalam konstruk yang terpisah, melainkan secara agregat dimana dimensi-dimensi individual dimasukkan untuk mendapatkan kualitas jasa secara keseluruhan [2].

Menurut Safira, Fardiana Nasbir bahwa Kepuasan pelanggan memberi beberapa manfaat bagi perusahaan sebagai berikut: [3]

1. Pelanggan yang puas akan menimbulkan komitmen yang kuat untuk membangun dan mempertahankan hubungan yang bernilai dalam jangka panjang.
2. Pelanggan yang puas akan menimbulkan perasaan pelanggan untuk mempercayai merek atau perusahaan. Pelanggan mempercayai merek atau perusahaan karena kinerja barang/jasa yang dibeli sesuai atau melebihi harapannya. Sebaliknya pelanggan tidak mempercayai merek atau perusahaan, karena kinerja barang/jasa tidak sesuai dengan harapannya.
3. Pelanggan yang puas meningkatkan kualitas hubungan antara pelanggan dengan perusahaan yang diukur dengan indikator. Hubungan itu penting, takut kehilangan hubungan dan perasaan bangga jika membangun hubungan dengan perusahaan atau merek tertentu.
4. Pelanggan yang puas akan menceritakan kebaikan dan keunggulan merek atau perusahaan kepada orang lain dan memberikan rekomendasi kepada orang lain untuk membeli barang/jasa yang dijual perusahaan yang sudah menjadi langganannya. Hal ini merupakan bentuk promosi yang efektif, karena cerita orang yang sudah merasakan barang/jasa yang dijual perusahaan mudah di percaya oleh calon konsumen.

Kualitas pelayanan adalah seberapa jauh perbedaan antara kenyataan dan harapan pelanggan atas layanan yang mereka terima. Kualitas pelayanan yaitu inti dari konsep kualitas layanan adalah menunjukkan segala bentuk aktualisasi kegiatan pelayanan yang memuaskan orang-orang yang menerima pelayanan sesuai dengan daya tanggap (*responsiveness*), menumbuhkan adanya jaminan (*assurance*), menunjukkan bukti fisik (*tangible*) yang dapat dilihatnya, menurut (*empathy*) dari orang-orang yang memberikan pelayanan sesuai dengan keandalannya (*reliability*) menjalankan tugas pelayanan yang diberikan secara konsekuen untuk memuaskan yang menerima pelayanan [4].

Menurut Gremler dan Brown dalam Hasan, bahwa loyalitas pelanggan adalah pelanggan yang tidak hanya membeli ulang suatu produk/jasa tetapi juga mempunyai komitmen dan sikap positif terhadap kepuasan jasa, misalnya merekomendasikan produk atau jasa ke orang lain untuk membeli [5].



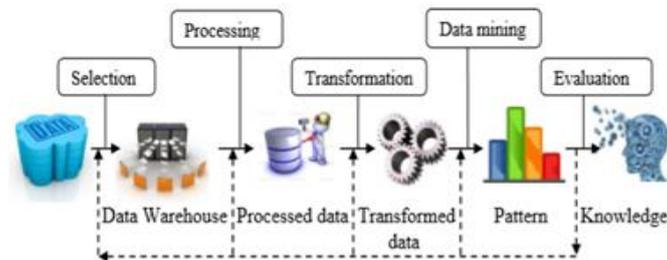
Gambar 1 Konsep Data Mining

Menurut Ayu Rizqi Oktaviana bahwa data mining adalah serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basis data. Informasi yang dihasilkan diperoleh dengan cara mengekstrasikan dan mengenali pola yang penting atau menarik dari data yang terdapat pada basis data. Data mining terutama digunakan untuk mencari pengetahuan yang terdapat dalam basis data yang besar sehingga sering disebut *Knowledge Discovery Databases (KDD)* [6].

Menurut Wulandari, Retno Tri Pohon keputusan atau *decision tree* adalah pemetaan, mengenai alternatif alternatif pemecahan masalah yang dapat diambil dari masalah tersebut. Pohon tersebut juga memperlihatkan faktor-faktor kemungkinan/probabilitas yang akan mempengaruhi alternatif-alternatif keputusan tersebut, disertai dengan estimasi hasil akhir yang akan didapat bila kita mengambil alternatif keputusan tersebut [7].

Menurut Jumhur, Annisak Izzaty bahwa *decision tree* merupakan metode klasifikasi yang paling sering digunakan. Karena dalam pengerjaannya tidak memerlukan waktu yang lama dan hasilnya pun mudah untuk dipahami dan banyak penelitian dalam kasus ini sering menggunakan *decision tree* untuk mendapatkan hasil yang maksimal [8].

Menurut Mohammad Rizki Ilham bahwa Data mining merupakan salah satu dari rangkaian Knowledge Discovery in Database (KDD) [9].



Gambar 2 Tahapan Data Mining

Menurut Larose, Ada beberapa tahapan dalam membuat sebuah pohon keputusan algoritma C4.5, yaitu [10],[11],[12]:

1. Mempersiapkan data training. Data training biasanya diambil dari data histori yang pernah terjadi sebelumnya atau disebut data masa lalu dan sudah dikelompokkan dalam kelas-kelas tertentu.
2. Menghitung akar dari pohon. Akar akan diambil dari atribut yang akan dipilih, dengan cara menghitung nilai gain dari masing-masing atribut, nilai gain yang paling tinggi yang akan menjadi akar pertama. Sebelum menghitung nilai gain dari atribut, hitung dulu nilai entropy.
3. Menghitung nilai gain menggunakan persamaan 2.
4. Ulangi langkah ke 2 dan ke 3 hingga semua record terpatisi.
5. Proses partisi pohon akan berhenti saat :
 - a. Semua record dalam simpul N mendapatkan kelas yang sama.
 - b. Tidak ada atribut yang dipartisi lagi.

Menurut suyanto, Secara istilah, entropy adalah keberbedaan atau keberagaman. Dalam data mining, entropy didefinisikan sebagai suatu parameter untuk mengukur heterogenitas (keberagaman) dalam suatu himpunan data. Semakin heterogen suatu himpuna data, semakin besar pula nilai entropy-nya. Secara matematis, entropy di rumuskan sebagai berikut:

$$Entropy(S) = \sum_i^c - p_i \log_2 p_i$$

Dimana c adalah jumlah nilai yang terdapat pada atribut target (jumlah target). Sedangkan p_i menyatakan porsi atau rasio antara jumlah sampel di kelas i dengan jumlah semua sampel pada himpunan data [12]. Secara istilah, information gain adalah perolehan informasi. Dalam data mining, information gain didefinisikan sebagai ukuran efektifitas suatu atribut dalam mengkalasifikasikan data. Secara matematis, information gain dari sudut atribut A, dituliskan sebagai berikut [13].

$$Gain(S, A) \equiv Entropy(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} Entropy(S_v)$$

dimana :

- A : Atribut
 V : Menyatakan suatu nilai yang mungkin untuk atribut a
 Values(A) : Himpunan nilai-nilai yang mungkin untuk atribut A
 |S_v| : Jumlah sampel untuk nilai v
 |S| : Jumlah seluruh sampel data
 Entropy(S_v) : Entropy untuk sampel-sampel yang memiliki nilai v

Metode Penelitian

Teknik pengumpulan data dalam penelitian ini yaitu kuesioner agar diperoleh data-data yang akurat, relevan, dan terpercaya [14]. Kuesioner merupakan suatu teknik pengumpulan data dengan cara memberikan seperangkat pertanyaan-pertanyaan tertulis kepada responden untuk dijawabnya [15]. Responden dalam penelitian ini yaitu pelanggan transportasi online (ojek online), kuesioner dibagikan kepada pelanggan-pelanggan yang sudah pernah memesan transportasi online (ojek online). Kuesioner terdiri dari 18 pertanyaan dan lima kategori dan jawaban-jawaban yang sudah disediakan oleh peneliti untuk dipilih oleh responden. Untuk atribut kuesioner terdiri dari harga transportasi online (ojek online), fasilitas yang diberikan oleh transportasi online (ojek online), pelayanan yang diberikan oleh driver transportasi online (ojek online), loyalitas pelanggan transportasi online (ojek online), dan hasil puas tidak puasnya pelanggan terhadap transportasi online (ojek online).

Untuk memperoleh data tentang variable dalam penelitian ini digunakan instrumen penelitian berupa kuesioner. Kuesioner adalah alat ukur yang terdiri dari sejumlah pertanyaan atau pernyataan tertulis yang harus dijawab atau diisi oleh responden [16]. Ada dua skala pengukuran yang dapat digunakan yaitu skala nominal dan interval.

Untuk menilai jawaban dari setiap pertanyaan dari kuesioner, dilakukan dengan menggunakan metode skala likert. Metode ini mengukur setiap penilaian responden melalui pemberian bobot setiap pertanyaan sebagai berikut:

Data yang digunakan pada penelitian ini berdasarkan survey dan teknik pengumpulan data menggunakan kuesioner yang diwujudkan dalam bentuk pertanyaan atau pernyataan yang diberikan kepada responden untuk diisi sesuai dengan keadaan sebenarnya. Metode yang digunakan untuk penelitian ini yaitu dengan menggunakan metode algoritma C4.5. Prosedur perhitungan penentuan kepuasan pelanggan transportasi online (ojek online).

Hasil dan Pembahasan

Data yang diambil dalam penelitian ini diperoleh dari pembagian kuesioner sebanyak 500 kuesioner. Data yang dikumpulkan yaitu data kuesioner yang dibagikan kepada pelanggan transportasi online (ojek online) dengan atribut nama, jenis kelamin, umur, harga, fasilitas, pelayanan, loyalitas, kepuasan.

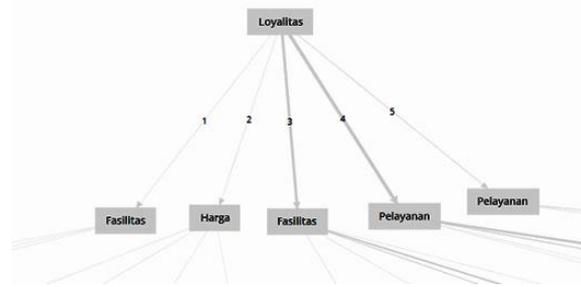
Dari pembagian 500 data kuesioner kepada pelanggan transportasi online (ojek online) diperoleh 425 kuesioner yang menyatakan "PUAS" dan 75 kuesioner yang menyatakan "TIDAK PUAS".

Tabel 2 Perhitungan Algoritma C4.5 Root Node

| Node | Attribute | Value | Label | | Jumlah Kasus | Entropy | Gain |
|------|--------------------|-------|-------|------------|--------------|---------|--------|
| | | | Puas | Tidak Puas | | | |
| 1 | Kepuasan Pelanggan | | 425 | 75 | 500 | 0,6098 | 0,0023 |
| | Harga | | | | | | |
| | 1 | 3 | 1 | 4 | 0,8113 | | |
| | 2 | 7 | 1 | 8 | 0,5436 | | |
| | 3 | 103 | 22 | 125 | 0,6712 | | |
| | 4 | 262 | 41 | 303 | 0,5718 | | |
| | 5 | 50 | 10 | 60 | 0,6500 | | |

| | | | | | |
|-----------|---|-----|----|-----|--------|
| Fasilitas | | | | | 0,0250 |
| | 1 | 5 | 2 | 13 | 0,9940 |
| | 2 | 11 | 4 | 19 | 0,9641 |
| | 3 | 144 | 30 | 172 | 0,5782 |
| | 4 | 236 | 33 | 240 | 0,5370 |
| | 5 | 29 | 6 | 56 | 0,6610 |
| Pelayanan | | | | | 0,0035 |
| | 1 | | 7 | 11 | 0,6194 |
| | 2 | | 7 | 24 | 0,7425 |
| | 3 | | 25 | 187 | 0,6677 |
| | 4 | | 29 | 238 | 0,5777 |
| | 5 | | 7 | 40 | 0,4912 |
| Loyalitas | | | | | 0,0270 |
| | 1 | | | | 0,9457 |
| | 2 | | | | 0,8709 |
| | 3 | | | | 0,5675 |
| | 4 | | | | 0,5347 |
| | 5 | | | | 0,6690 |

Hasil Entropy total adalah 0,6098 dengan jumlah kasus 500 data yaitu 425 “PUAS” dan 75 “TIDAK PUAS”. Dari hasil perhitungan diatas, dipilih atribut dengan nilai information gain sebagai nilai terbesar adalah 0,0270 yaitu atribut loyalitas Selanjutnya atribut loyalitas digunakan untuk menjadi akar (root). Pada sub-node selanjutnya tidak dapat digunakan lagi sebagai ekspansi tree. Berikut bentuk pohon keputusan root node. Berikut hasil pengujian decision tree menggunakan rapidminer



Gambar 3 Pohon Keputusan Bagian Akar

Pada gambar 3. merupakan pohon keputusan bagian simpul paling atas atau node akar (root node). Atribut yang gainnya terbesar menjadi akar adalah atribut harga . Atribut harga menghasilkan 4 node yaitu 1, 2, 3, 4, dan 5. Pada internal node 1 menghasilkan node fasilitas, internal node 2 menghasilkan node harga, internal node 3 menghasilkan fasilitas, internal node 4 menghasilkan node pelayanan, dan internal node 5 menghasilkan node pelayanan. Dari perhitungan tabel 4.5 diperoleh akar (root node) dari pohon adalah loyalitas dan menghasilkan 5 node yaitu 1, 2, 3, 4, dan 5.

Tabel 3 Perhitungan Node 1

| Perhitungan Node 1 | Jumlah | Puas | Tidak Puas | Entropy | Gain |
|--------------------|--------|------|------------|---------|--------|
| Loyalitas | 1 | 11 | 4 | 7 | 0,9457 |
| Harga | | | | | 0,2630 |
| | 1 | 1 | 0 | 1 | 0,0000 |
| | 2 | 1 | 0 | 1 | 0,0000 |
| | 3 | 2 | 1 | 1 | 1,0000 |
| | 4 | 6 | 2 | 4 | 0,9183 |
| | 5 | 1 | 1 | 0 | 0,0000 |
| Fasilitas | | | | | 0,3316 |
| | 1 | 4 | 0 | 4 | 0,0000 |
| | 3 | 3 | 2 | 1 | 0,9183 |
| | 4 | 4 | 2 | 2 | 1,0000 |
| Pelayanan | | | | | 0,3225 |
| | 1 | 2 | 1 | 1 | 1,0000 |
| | 3 | 2 | 0 | 2 | 0,0000 |
| | 4 | 5 | 3 | 2 | 0,9710 |
| | 5 | 2 | 0 | 2 | 0,0000 |

Pada perhitungan internal node loyalitas “1” gain tertinggi adalah atribut fasilitas dan mempunyai 3 node, node 1 menghasilkan keputusan “TIDAK PUAS”, sedangkan node 3 dan 4 menghasilkan node.

Tabel 4 Perhitungan Node 1.1

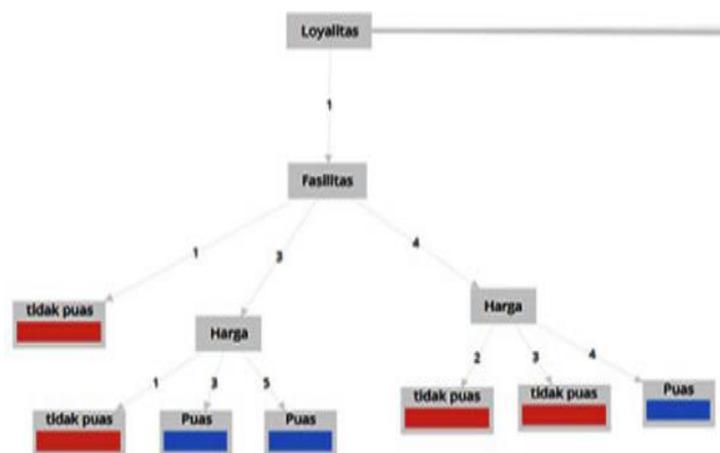
| Perhitungan Node 1.1 | Jumlah | Puas | Tidak Puas | Entropy | Gain |
|----------------------|--------|------|------------|---------|--------|
| Fasilitas | 3 | 3 | 2 | 1 | 0,9183 |
| Harga | | | | | 0,9183 |
| 1 | 1 | 0 | 1 | 0,0000 | |
| 3 | 1 | 1 | 0 | 0,0000 | |
| 5 | 1 | 1 | 0 | 0,0000 | |
| Pelayanan | | | | | 0,9183 |

Pada perhitungan node fasilitas “3” gain tertinggi adalah atribut harga dan mempunyai 3 node, node 1 menghasilkan keputusan “TIDAK PUAS”, node 3 menghasilkan keputusan “PUAS”, sedangkan node 4 menghasilkan keputusan “PUAS”.

Table 5 Perhitungan Node 1.2

| Perhitungan Node 1.2 | Jumlah | Puas | Tidak Puas | Entropy | Gain |
|----------------------|--------|------|------------|---------|--------|
| Fasilitas | 4 | 4 | 2 | 2 | 1,0000 |
| Harga | | | | | 1,0000 |
| 2 | 1 | 0 | 1 | 0,0000 | |
| 3 | 1 | 0 | 1 | 0,0000 | |
| 4 | 2 | 2 | 0 | 0,0000 | |
| Pelayanan | | | | | 0,0000 |
| 4 | 4 | 2 | 2 | 1,0000 | |

Pada perhitungan node fasilitas “4” gain tertinggi adalah atribut harga dan mempunyai 3 node, node 2 menghasilkan keputusan “TIDAK PUAS”, node 3 menghasilkan keputusan “TIDAK PUAS”, sedangkan node 4 menghasilkan keputusan “PUAS”. Berikut hasil pengujian decision tree node 1, 1.1, dan 1.2 menggunakan rapidminer



Gambar 4 Decision tree internal node 1

Pada internal node 1 yaitu node fasilitas memiliki 3 node 1, 3, 4 :

1. Node 1 menghasilkan keputusan “PUAS”.
2. Node 3 pada node harga menghasilkan 3 node yaitu 1, 3, dan 5.
 - a. Node 1 menghasilkan keputusan “TIDAK PUAS”.
 - b. Node 3 menghasilkan keputusan “PUAS”.
 - c. Node 5 menghasilkan keputusan “PUAS”.
3. Sedangkan pada node 4 yaitu node harga menghasilkan 3 node yaitu 2, 3, dan 4.
 - a. Node 2 menghasilkan keputusan “TIDAK PUAS”.
 - b. Node 3 menghasilkan keputusan “TIDAK PUAS”.
 - c. Node 4 menghasilkan keputusan “PUAS”.

Tabel 6 Perhitungan Node 2

| Perhitungan Node 2 | Jumlah | Puas | Tidak Puas | Entropy | Gain |
|--------------------|--------|------|------------|---------|--------|
| Loyalitas | 2 | 24 | 17 | 7 | 0,8709 |
| Harga | | | | | 0,1786 |
| | 2 | 1 | 1 | 0 | 0,0000 |
| | 3 | 12 | 9 | 3 | 0,8113 |
| | 4 | 9 | 7 | 2 | 0,7642 |
| | 5 | 2 | 0 | 2 | 0,0000 |
| Fasilitas | | | | | 0,1600 |
| | 2 | 4 | 2 | 2 | 1,0000 |
| | 3 | 9 | 5 | 4 | 0,9911 |
| | 4 | 4 | 4 | 0 | 0,0000 |
| | 5 | 7 | 6 | 1 | 0,5917 |
| Pelayanan | | | | | 0,1501 |
| | 2 | 2 | 1 | 1 | 1,0000 |
| | 3 | 11 | 8 | 3 | 0,8454 |
| | 4 | 6 | 3 | 3 | 1,0000 |
| | 5 | 5 | 5 | 0 | 0,0000 |

Pada perhitungan internal node loyalitas “2” gain tertinggi adalah atribut harga dan mempunyai 4 node, node 2 menghasilkan node keputusan “PUAS”, node 3 dan 4 menghasilkan node, sedangkan node 5 menghasilkan node keputusan “TIDAK PUAS”

Table 7 Perhitungan Node 2.1

| Perhitungan Node 2.1 | Jumlah | Puas | Tidak Puas | Entropy | Gain |
|----------------------|--------|------|------------|---------|--------|
| Harga | 3 | 12 | 9 | 3 | 0,8113 |
| Fasilitas | | | | | 0,0782 |
| | 2 | 3 | 2 | 1 | 0,9183 |
| | 3 | 7 | 5 | 2 | 0,8631 |
| | 4 | 2 | 2 | 0 | 0,0000 |
| Pelayanan | | | | | 0,0430 |
| | 3 | 10 | 8 | 2 | 0,7219 |
| | 4 | 2 | 1 | 1 | 1,0000 |

Pada perhitungan node harga “3” gain tertinggi adalah atribut fasilitas dan mempunyai 3 node, node 2 menghasilkan node, node 3 menghasilkan 5 (puas) dan 2 (tidak puas) dengan hasil keputusan “PUAS” 5(puas), sedangkan node 4 menghasilkan node keputusan “PUAS”.

Table 8 Perhitungan Node 2.1.1

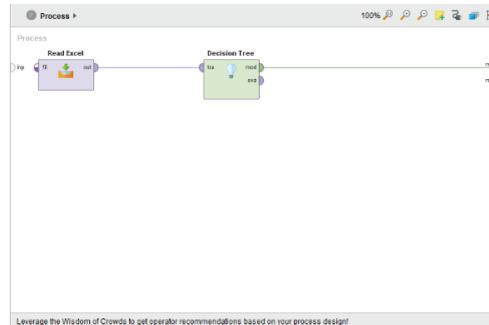
| Perhitungan Node 2.1 | Jumlah | Puas | Tidak Puas | Entropy | Gain |
|----------------------|--------|------|------------|---------|--------|
| Fasilitas | 2 | 3 | 2 | 1 | 0,9710 |
| Pelayanan | | | | | 0,9710 |
| | 3 | 2 | 2 | 0 | 0,0000 |
| | 4 | 1 | 0 | 1 | 0,0000 |

Pada perhitungan node fasilitas “2” gain tertinggi adalah atribut pelayanan dan mempunyai 2 node keputusan, node 3 menghasilkan node keputusan “PUAS”, node 4 menghasilkan node keputusan “TIDAK PUAS”.

Table 9 Perhitungan Node 2.2

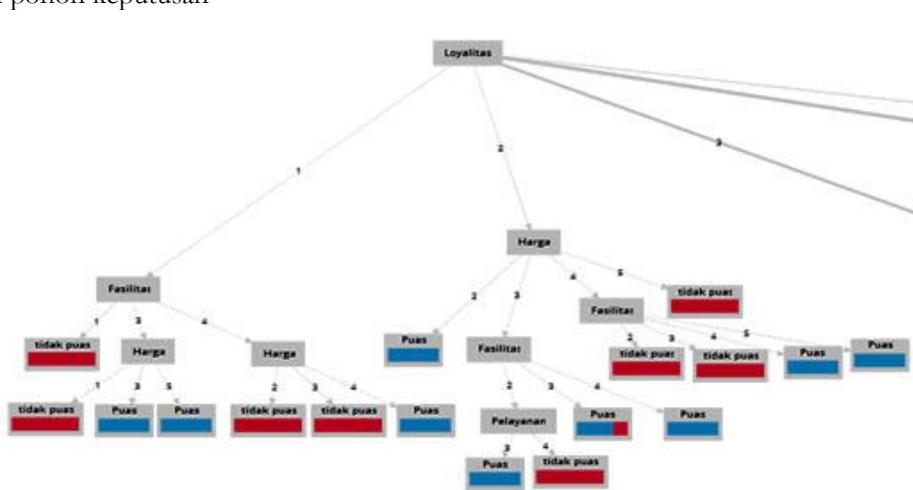
| Perhitungan Node 2.2 | Jumlah | Puas | Tidak Puas | Entropy | Gain |
|----------------------|--------|------|------------|---------|--------|
| Harga | 4 | 9 | 7 | 2 | 0,7642 |
| Fasilitas | | | | | 0,7642 |
| | 2 | 1 | 0 | 1 | 0,0000 |
| | 3 | 1 | 0 | 1 | 0,0000 |
| | 4 | 1 | 1 | 0 | 0,0000 |
| | 5 | 6 | 6 | 0 | 0,0000 |
| Pelayanan | | | | | 0,3198 |
| | 4 | 4 | 2 | 2 | 1,0000 |
| | 5 | 5 | 5 | 0 | 0,0000 |

Tahap selanjutnya data akan di proses menggunakan bantuan tools rapidminer dan untuk pengujianya menggunakan split validation yang ada didalam rapidminer, maka data akan dilihat nilai accuracy, pohon keputusan, dan rulenya. Pada tahap pengujian model ini data yang digunakan telah melewati tahap preprosesing. Selanjutnya data akan dilihat nilai accuracy, pohon keputusan, dan rule pada rapidminer.

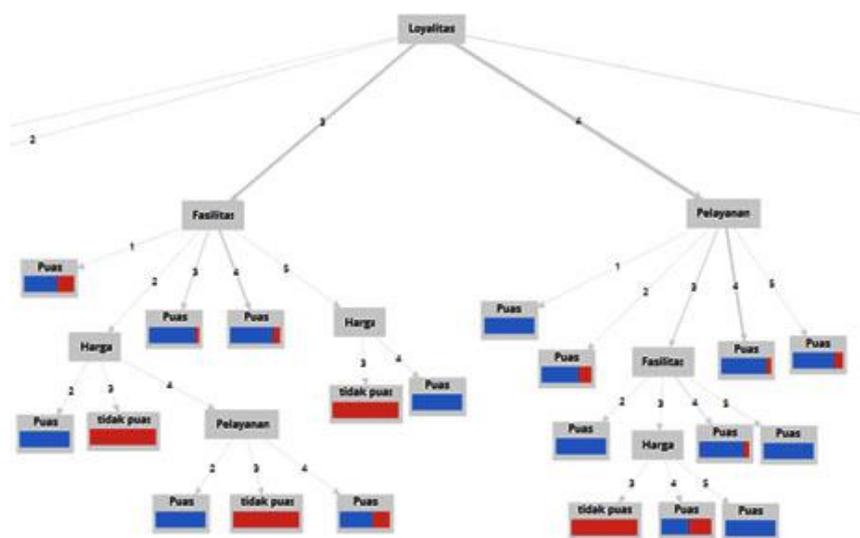


Gambar 5 Proses pengolahan data pada RapidMiner

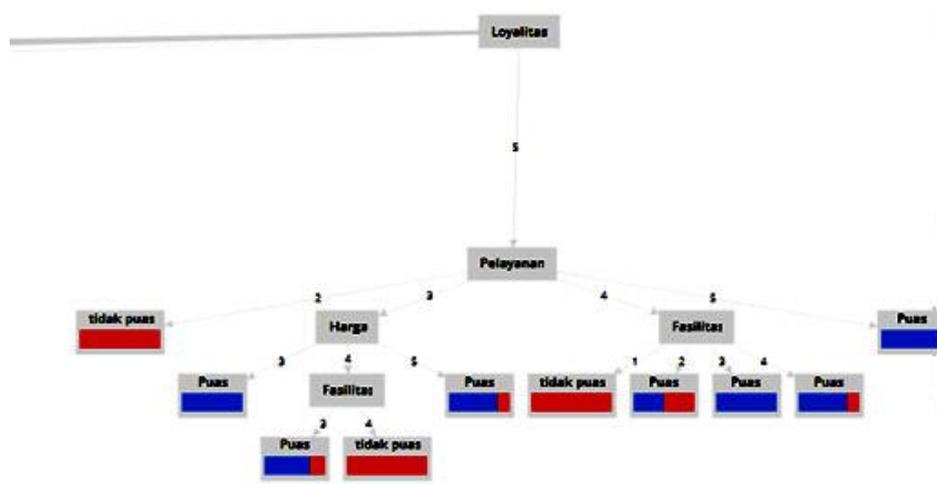
Pada model pengujian dibuat dengan read excel digunakan untuk import data kepuasan pelanggan transportasi online (ojek online) dihubungkan dengan operator decision Tree yang digunakan untuk menampilkan pohon keputusan



Gambar 6 Pohon Keputusan bagian 1

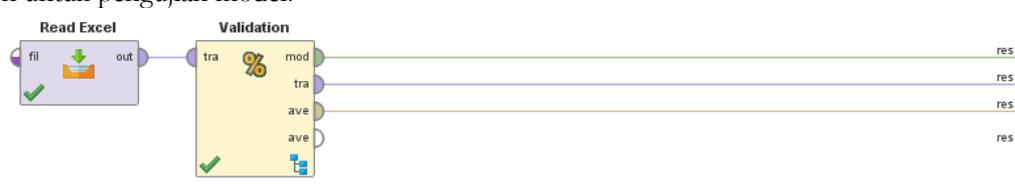


Gambar 7 Pohon Keputusan bagian 2



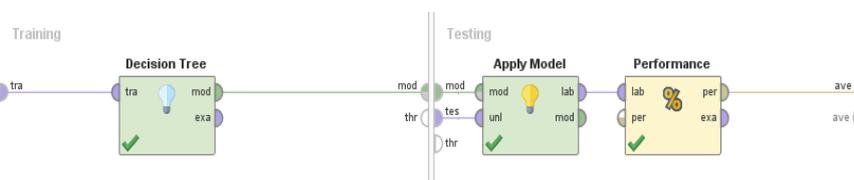
Gambar 8 Pohon Keputusan bagian 3

Pengujian ini digunakan 500 data untuk mengetahui nilai accuary, precison, dan recall digunakan split validation untuk pengujian model.



Gambar 9 Pengujian Algoritma Pada Rapidminer

dataset kepuasan pelanggan transportasi online (ojek online) dihubungkan dengan operator split validation yang didalamnya terdapat proses seperti gambar 10 dibawah ini.



Gambar 10 Pengujian menggunakan split validatin

accuracy: 83.22%

| | true Puas | true tidak puas | class precision |
|------------------|-----------|-----------------|-----------------|
| pred. Puas | 121 | 19 | 86.43% |
| pred. tidak puas | 6 | 3 | 33.33% |
| class recall | 95.28% | 13.64% | |

Gambar 11 Hasil Accurasi

Hasil accuracy menghasilkan 83,22% dapat dilihat pada gambar diatas 4.20. Jumlah true positive (tp) sebanyak 121 record, false positive (fp) sebanyak 6 record, jumlah true negative (tn) sebanyak 3 record dan jumlah false negative (fn) sebanyak 19 record. Accuracy yang dihasilkan dari pengujian pada rapidminer sebesar 83,22%. Berikut adalah perhitungan accuracy:

$$\begin{aligned}
 \text{Accuracy} &= \frac{tp + tn}{tp + tn + fp + fn} * 100\% \\
 &= \frac{121 + 3}{121 + 3 + 6 + 19} * 100\% \\
 &= \left(\frac{124}{149}\right) * 100\% \\
 &= 0,832 * 100\% \\
 &= 83,22\%
 \end{aligned}$$

precision: 33.33% (positive class: tidak puas)

| | true Puas | true tidak puas | class precision |
|------------------|-----------|-----------------|-----------------|
| pred. Puas | 121 | 19 | 86.43% |
| pred. tidak puas | 6 | 3 | 33.33% |
| class recall | 95.28% | 13.64% | |

Gambar 12 Hasil precision

$$\begin{aligned}
 Precision &= \frac{tp}{tp + fn} * 100\% \\
 &= \frac{121}{121 + 19} * 100\% \\
 &= \frac{121}{140} * 100\% \\
 &= 86,43\%
 \end{aligned}$$

Nilai recall dihitung dengan cara membagi jumlah data benar yang bernilai positif (true positif) dibagi dengan jumlah data besar yang bernilai positif (true positif) dan data salah yang bernilai (false positif).

recall: 13.64% (positive class: tidak puas)

| | true Puas | true tidak puas | class precision |
|------------------|-----------|-----------------|-----------------|
| pred. Puas | 121 | 19 | 86.43% |
| pred. tidak puas | 6 | 3 | 33.33% |
| class recall | 95.28% | 13.64% | |

Gambar 13 Hasil Reccal

$$\begin{aligned}
 Recall &= \frac{tp}{tp + fn} * 100\% \\
 &= \frac{121}{121 + 6} * 100\% \\
 &= \frac{121}{127} * 100\% \\
 &= 0,9527 * 100\% \\
 &= 95,28\%
 \end{aligned}$$

Hasil pengujian menghasilkan accuracy, precision, dan recall sebagai berikut :

Tabel 10 Hasil Pengujian

| Accuracy | Precision | Recall |
|-----------------|------------------|---------------|
| 83,22% | 86,43% | 95,28% |

Berdasarkan table yang telah dijelaskan diatas, maka dapat diketahui bahwa nilai accuracy 83,22%, precision 86,43%, dan recall 95,28%.

Kesimpulan

Dari hasil penelitian yang telah dilakukan oleh peneliti, maka dapat disimpulkan bahwa kepuasan pelanggan transportasi online (ojek online) dapat diprediksi dan dievaluasi dengan memanfaatkan teknik data mining menggunakan algoritma C4.5 untuk memprediksi dari kepuasan pelanggan serta hasil percobaan dan pengujian prediksi kepuasan pelanggan transportasi online (ojek online) dengan RapidMiner menggunakan metode decision tree C4.5, diperoleh akurasi 83,22% dengan kriteria good classification menggunakan confusion matrix.

Daftar Rujukan

- [1] Jamhur, Annisak Izzaty. (2016). "Penerapan Data Mining Untuk Menganalisa Jumlah Pelanggan Aktif Dengan Menggunakan Algoritma C4.5". *Journal*, Volume 23, No. 3.
- [2] Harpadeles & Jushermi, n.d. (2016). "Pengaruh Kualitas Pelayanan Dan Nilai Pelanggan Terhadap Kepuasan Dan Loyalitas Pelanggan Trans Metro Pekanbaru". *Journal, JOM FEKON* Vol. 3 No. 1.
- [3] Khusaini, Ahmad. (2016). "Analisis Kualitas Pelayanan Terhadap Kepuasan Konsumen Di SPA CLUB Arena Yogyakarta", Tugas Akhir Fakultas Ilmu Keolahragawan. *Yogyakarta : Universitas Negeri Yogyakarta*.
- [4] Maris, Eki Ruziqa. (2016). "Analisis Kepuasan Pelanggan Menggunakan Algoritma C4.5". *Journal*, eprints.dinus.ac.id/16479/1/jurnal_15407.

- [5] Nasbir, Safira Farizah. (2017). “Pengaruh Kualitas Pelayanan Terhadap Kepuasan Pelanggan Salon Hair Nets” *Tugas Akhir Ekonomi Dan Bisnis. Kendari.*
- [6] Oktaviana, Ayu Rizki. (2016). “Penerapan Data Mining Klasifikasi Pola Nasaba Menggunakan Algoritma C4.5 Pada Bank BRI Batang”.
- [7] Panjaitan, Januar Efendi., & Yulianti, Ai Lili. (2016). “Pengaruh Kualitas Pelayanan Terhadap Kepuasan Pelanggan Pada JNE Cabang Bandung”. *Journal Of DeReMa Jurnal Manajemen*, Volume 11. No.2.
- [8] Rohman, Ibnu Fachtur. (2016). “Penerapan Algoritma C4.5 Pada Kepuasan Pelanggan Perum DAMRI”.
- [9] S, Mochammad Rizki Ilham. (2016). “Implementasi Data Mining Menggunakan Algoritma C4.5 Untuk Prediksi Kepuasan Pelanggan Taksi KOSTI”.
- [10] Larose, Daniel, *Data Mining Methods and Models. New Jersey: John Welby & Sons, Inc. Hoboken*
- [11] Romli, I., Kharida, F., & Naya, C. (2020). Penentuan Kepuasan Pelanggan Terhadap Pelayanan Kantor Pelayanan Pajak Menggunakan C4. 5 dan PSO. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 4(2), 296-302.
- [12] Fatchan, M., Afriantoro, I., Sari, P. A., & Kodratilah, E. Y. (2021). Pengujian Akurasi Data Potensi Kepuasan Pelanggan Kereta Commuterline (KRL) Dengan Algoritma C4. 5. *Pelita Teknologi*, 16(1), 8-14.
- [13] Suyanto. (2019). Data Mining Untuk Klasifikasai dan Klasterisasi Data. *Informatika Bandung.*
- [14] Eryanti, N. F. (2017). Pengaruh Model Pembelajaran Kooperatif Terhadap Hasil Belajar Fisika Aspek Kognitif dan Afektif Peserta Didik Kelas X SMA N 2 Bantul.
- [15] Febriarini, A. S., & Astuti, E. Z. (2019). Penerapan Algoritma C4. 5 untuk Prediksi Kepuasan Penumpang Bus Rapid Transit (BRT) Trans Semarang. *Jurnal Eksplora Informatika*, 8(2), 95-103.
- [16] Piliandani, N. M. I., Pradnyanitasari, P. D., & Saputra, K. A. K. (2020). Pengaruh persepsi dan pengetahuan akuntansi pelaku usaha mikro kecil dan menengah terhadap penggunaan informasi akuntansi. *Jurnal Akuntansi, Ekonomi Dan Manajemen Bisnis*, 8(1), 67-73.
- [17] Aditiya, w. (2021). Pengaruh Layanan Bimbingan Kelompok Teknik Self Management Terhadap Kedisiplinan Beribadah Remaja Di Desa Banjarrejo 38b Batanghari Lampung Timur (*doctoral dissertation, <https://ummetro.ac.id/>*).