

# DECISION TREE PADA PENILAIAN PERFORMA LAYANAN DRIVER GOJEK DENGAN ALGORITMA C4.5

Frederikus Cyriaco Leu Le Uape<sup>1</sup>, Erlin Elisa<sup>2</sup>

<sup>1</sup>Mahasiswa Program Studi Sistem Informasi, Universitas Putera Batam

<sup>2</sup>Dosen Program Studi Sistem Informasi, Universitas Putera Batam

email: [pb191510051@upbatam.ac.id](mailto:pb191510051@upbatam.ac.id)

## ABSTRACT

*Transportation is an important means to support daily activities and operations, be it private transportation or public transportation. This study will analyze the level of customer satisfaction as a public vehicle user, in this case the author utilizes data on customers who use and utilize the gojek application as a means of transportation, using a questionnaire data collection technique, the writer will determine client satisfaction levels using the variables Tangible, Reliability, Responsiveness, Assurance, Empathy. The analytical Data Mining with the is a technique employed in this investigation. concept of the Algorithm C4.5 classification which can show the results of the highest Entrophy and gain values modeled as a decision tree with the aim that the results obtained can be a guide for gojek partners to service and increase customers. Considering the outcomes of the classicization It is demonstrated by the C4.5 algorithm that the accuracy achieves 75.33% where the factors that become the assessment of satisfaction are Tangible, Reliability.*

**Keywords :** *Data Mining, Decision Tree, C4.5 Algorithm, Online Transportation, and Customer Satisfaction*

## PENDAHULUAN

Transportasi ke tempat kerja, sekolah, universitas, pasar, transportasi, dll. Hal tersebut merupakan salah satu tuntutan masyarakat saat ini karena dapat memudahkan segala tugas sehari-hari. Berdasarkan perkembangan penggunaan mobil di Indonesia pada tahun 2020, 51,24% tenaga kerja Indonesia menggunakan mobil pribadi atau perusahaan. Sedangkan jumlah pengguna angkutan umum masih rendah yaitu 41,93%. Bagian ini telah berubah sejak tahun lalu. Pada tahun 2019, pengguna atau *user* mobil pribadi menyumbang 83,76% kendaraan umum 11,81%. Meskipun demikian kendaraan

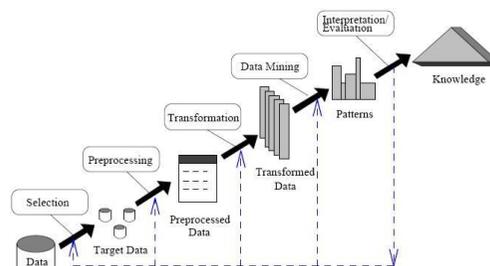
umum juga menjadi pilihan bagi Angkutan umum ini digunakan oleh masyarakat yang belum atau sudah memiliki mobil. Ada banyak orang yang bekerja di Gojek Batam. Gojek merupakan sebuah perusahaan swasta yang merupakan pusat bisnis yang fokus pada bisnisnya yaitu jasa transportasi yaitu layanan motoxi online. Gojek bekerja sama dengan *driver* berpengalaman untuk pengiriman, pemesanan makanan, belanja, dan perjalanan sebagai solusi terbaik. Untuk memuaskan pelanggan dan tetap berada di depan banyak pesaing saat ini di dunia pengiriman online, Gojek harus terus meningkatkan layanan pelanggannya. Untuk itu, harus

kepuasan konsumen diukur dari sudut pandang pengguna layanan *online* sehingga *driver* Gojek dapat belajar dan mengevaluasi diri mereka sendiri dan bahwa mereka memberikan layanan yang baik hingga saat ini. Menurut perjalanan penulis hingga saat ini akan tetapi data yang telah terkumpul selama ini ada di gudang data gojek sendiri, *driver* hanya mengetahui penilaian masing-masing, karena setiap *driver* memiliki pengalamannya masing-masing dalam melakukan pekerjaannya, Maka untuk mendapatkan contoh *review* pelanggan yang lengkap untuk meningkatkan citra kerja *driver*, penting untuk memeriksa peringkat pelanggan. dan semua orang senang benar-benar menjadi faktor utama tertariknya orang menjadi langganan. Dalam penelitian ini peneliti menerapkan Salah satu kelebihan algoritma C4.5 adalah metode pohon keputusan. Kelebihan algoritma ini adalah bahwa itu dapat membuat pohon keputusan yang mudah dipahami, memiliki tingkat akurasi yang dapat diterima, dan mampu menangani data numerik dan diskret. (Adriansa dkk., 2022).

### KAJIAN TEORI

#### 2.1 KDD (Knowledge Discovery in Database)

*KDD adalah tindakan yang memiliki banyak tujuan dan tahapan dalam pengumpulan data yang besar. Tahapan berikut adalah tahapan dari Knowledge Discovery dalam database dengan detail* (T. A. Q. Putri dkk., 2023).



**Gambar 1.** Tahapan *Knowledge Discovery in Database (KDD)*

#### 2.2 Data Mining

Data mining terutama digunakan untuk menemukan pengetahuan dalam basis data yang besar, sehingga disebut Knowledge Discovery in Databases (KDD). Data mining menghasilkan nilai tambah dengan mengekstrak dan mengidentifikasi pola penting atau menarik dari data yang terdapat dalam basis data. (Pasaribu, 2021).

#### 2.3 Klasifikasi

Metode yang dikenal sebagai klasifikasi bertujuan untuk mengelompokkan kumpulan data secara sistematis menurut aturan dan standar yang telah ditetapkan sebelumnya. Menurut Magfirah (2019), klasifikasi adalah salah satu metode data mining yang digunakan untuk menentukan kelas objek yang belum diketahui kelasnya. Pada tahap awal klasifikasi, proses pelatihan dan pengujian akan dilakukan dengan dataset yang telah diketahui kelas objeknya. Empat komponen utama membentuk proses klasifikasi: (Zahida & Sulistiyoningrum, 2020).

#### 2.4 Algoritma C4.5

Beberapa algoritma, seperti ID3, CART, dan C.45, dapat digunakan untuk membuat pohon keputusan. pengembangan algoritma ID3. Pohon

keputusan yang biasanya dibangun dengan algoritma C4.5 adalah sebagai berikut (Mashlahah, 2020)..

- Pilih atribut sebagai akar
- Buat cabang untuk tiap nilai
- Beri kasus pada cabang
- Ulangi prosedur untuk setiap cabang untuk memastikan semua kasus pada cabang memiliki kelas yang sama.

Nilai Gain tertinggi dari atribut yang ada merupakan dasar untuk memilih atribut sebagai akar. Persamaan berikut digunakan untuk menghitung Gain.

(Sutabri, 2019) :

$$Gain(S, A) = Entropy(S) - \sum_i^n \frac{|s_i|}{|S|} * Entropy(s_i)$$

**Rumus 1.** Menghitung nilai gain

Setelah memperoleh nilai gain, perhitungan lain yang perlu dilakukan adalah menemukan nilai entropy. Nilai entropy digunakan untuk menentukan seberapa informatif sebuah atribut yang dimasukkan untuk menghasilkan atribut yang dikeluarkan. Rumus dasar entropy adalah sebagai berikut.

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

**Rumus 2.** Menghitung nilai entropy

### 2.5 Decision Tree

Salah satu teknik klasifikasi yang efektif dan terkenal adalah pohon keputusan. Metode pohon keputusan mengubah fakta yang signifikan menjadi pohon keputusan yang menggambarkan aturan yang mudah diinterpretasikan oleh manusia. Selain itu, pohon keputusan sangat membantu dalam analisis data karena mereka dapat

mengidentifikasi hubungan tersembunyi antara sebuah variabel target dan banyak variabel input (Shiddiq dkk., 2019). Model pohon keputusan digunakan untuk mengurangi populasi yang heterogen menjadi lebih kecil (homogen) dengan mempertimbangkan variabel tujuan.

### 2.6 Rapid Miner

*RapidMiner adalah perangkat lunak open-source. RapidMiner menawarkan analisis data mining, text mining, dan prediksi. Berbagai teknik deskriptif dan prediksi membantu pengguna membuat keputusan yang lebih baik. Lebih dari lima ratus operator data mining tersedia di RapidMiner, termasuk operator untuk input, output, data preprocessing, dan visualisasi. RapidMiner adalah program analisis data yang berdiri sendiri dan juga dapat diintegrasikan pada produknya sendiri. RapidMiner berjalan di semua sistem operasi karena dibuat dengan bahasa Java.*

### 2.7 Kepuasan Pelanggan

Kepuasan pelanggan dapat didefinisikan sebagai ketika harapan pelanggan terhadap pelayanan yang diberikan sesuai dengan layanan yang diberikan secara langsung kepada pelanggan. Jika layanan penyedia suatu jasa jauh dari harapan pelanggan, hal tersebut menyebabkan pelanggan akan merasa kecewa. Namun, pelanggan akan merasa puas jika layanan memenuhi harapan mereka. Pengalaman pelanggan saat menggunakan layanan dapat menentukan harapan mereka. penyedia layanan, dengan kata-kata orang lain, maupun dengan iklan (T. A. Q. Putri dkk., 2023).

### 2.8 Kualitas Pelayanan

Kualitas layanan adalah seberapa jauh perbedaan antara kenyataan dan harapan pelanggan tentang layanan yang mereka terima. Kualitas layanan adalah inti dari konsep kualitas layanan dan mencakup segala bentuk aktualisasi kegiatan layanan yang memuaskan orang-orang yang menerima layanan sesuai dengan daya tanggap (responsiveness), menumbuhkan adanya jaminan (assurance), menunjukkan bukti fisik (tangible) yang dapat dilihatnya, menurut (empathy) dari orang-orang yang menerima layanan. (Magfirah, 2019).

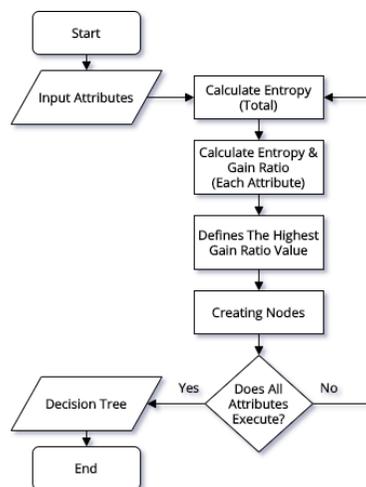
### METODE PENELITIAN

Pada Penelitian ini, penulis akan menggunakan metode *data mining* algoritma C4.5, Adapun langkah kerja dalam penelitian ini dapat di gambarkan pada alur berikut.



**Gambar 2.** Desain penelitian (Sumber: Data penelitian, 2023)

Algoritma C4.5 merupakan algoritma klasifikasi yang menghitung berdasarkan nilai entropy dan gain tertinggi, adapun model dari algoritma ini adalah:



**Gambar 3.** Flowchart model penelitian algoritma C4.5

### HASIL DAN PEMBAHASAN

Data mining menggunakan decision tree dan Algoritma C4.5 untuk menganalisis data penilaian performa layanan driver Gojek. Data kuesioner adalah sumber dari penelitian ini. Data dikumpulkan dengan membagi 133 kuesioner kepada pelanggan transportasi online Go-Jek dengan informasi seperti nama, jenis kelamin, umur, alamat, responsifitas, keyakinan, nyata, empati, dan kepercayaan. 126 kuesioner menyatakan "PUAS", dan 7 kuesioner menyatakan "TIDAK PUAS".

Berikut ini adalah tahap pohon keputusan (pohon keputusan) dengan algoritma. C4.5 menggunakan 133 data pelanggan tentang kepuasan dan ketidakpuasan transportasi Go-Jek; 126 "PUAS" dan 7 "TIDAK PUAS" dihasilkan. Responsif, yakin, nyata, kepercayaan, dan empati

adalah atribut yang digunakan. Perhitungan algoritma C4.5 dapat ditemukan di sini.

Metode algoritma C4.5, berdasarkan atribut yang sudah ada, akan digunakan untuk menghitung nilai akar berdasarkan masalah yang ada. Perhitungan awal akan mencari nilai akar dari setiap atribut dengan membandingkan gain tertinggi, dan sebelum mencari gain tersebut, perlu menghitung entropy untuk mengetahui hasil dari

gain masing-masing atribut. Berdasarkan masalah yang ada, metode algoritma C4.5 akan digunakan untuk menghitung nilai akar berdasarkan atribut yang sudah ada. Perhitungan awal akan mencari nilai akar dari setiap atribut dengan membandingkan gain tertinggi, tetapi sebelum mencari gain tersebut, harus menghitung entrophynya.

**Tabel 1.** Jumlah kasus

Jumlah Kasus	Puas	Tidak Puas
133	126	7

Sumber: (Peneliti, 2023)

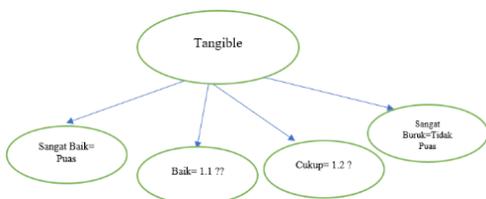
**Tabel 2.** Hasil perhitungan

Node	Kriteria	Value	Jumlah Kasus	Puas	Tidak Puas	Entropy	Gain
1.1	Total		133	126	7	0,999456	
1	Responsive						0,0583726
		Sangat Baik	48	46	2	0,720787	
		Baik	68	68	0	0,893635	
		Cukup	13	12	1	0,593387	
		Buruk	2	0	2	0	
		Sangat Buruk	2	0	2	0	
2	Assurance						0,738482
		Sangat Baik	50	48	2	0,7603645	
		Baik	73	70	3	0,7182916	
		Cukup	10	8	2	0,3472514	
		Buruk	0	0	0	0	
		Sangat Buruk	0	0	0	0	
3	Tangible						0,7399264
		Sangat Baik	32	30	2	0,627252	
		Baik	88	86	2	0,829735	
		Cukup	13	10	3	0,528365	
		Buruk	0	0	0	0	
		Sangat Buruk	0	0	0	0	

<b>4</b>	<b>Reliability</b>					<b>0,623447</b>
	Sangat Baik	30	30	0	0,512725	
	Baik	68	66	2	0,723625	
	Cukup	33	30	3	0,518237	
	Buruk	2	0	2	0	
	Sangat Buruk	0	0	0	0	
<b>5</b>	<b>Empathy</b>					<b>0,687447</b>
	Sangat Baik	32	30	2	0,526254	
	Baik	69	68	1	0,718362	
	Cukup	32	28	4	0,472527	
	Buruk	0	0	0	0	
	Sangat Buruk	0	0	0	0	

Sumber: (Peneliti, 2023)

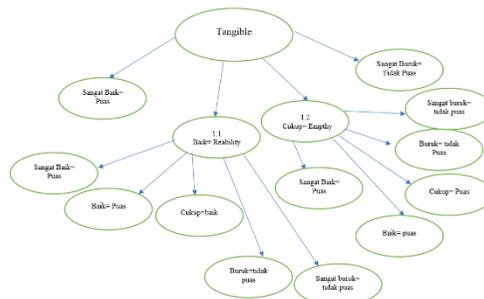
Dengan nilai information gain terbesar 0,7399264, atribut tangible digunakan sebagai akar. Tidak dapat digunakan lagi sebagai ekspansi pohon pada sub-node berikutnya. Bentuk pohon keputusan root node seperti ini.



**Gambar 4.** Pohon keputusan node 1

Gambar di atas menunjukkan bahwa akar yang sudah ditentukan pada perhitungan di atas tangible, yang berarti bahwa jumlah kasus atribut tangible sangat baik adalah 2 untuk "Tidak Puas" dan 30 untuk "Puas", yang berarti bahwa setiap kasus dengan nilai sangat baik hasilnya adalah Puas. Jumlah kasus atribut tangible sangat buruk adalah 0 untuk "Tidak Puas" dan 0 untuk "Puas", yang berarti bahwa setiap Jika begitu menemukan nilai akar, kemudian cari Node 1.1 dan Node 1.2

untuk mengetahui cabang dari akar tersebut. Untuk mempermudah perhitungan, cari Node 1.1 terlebih dahulu dengan menggunakan rumus yang sama dengan yang digunakan untuk menemukan entropy dan keuntungan tertinggi dari data filter tangible "baik".



**Gambar 5.** Pohon keputusan node 3 Hasil analisis data dihasilkan dengan menggunakan Microsoft Excel dan metode algoritma C4.5. Berdasarkan data yang digunakan untuk menghitung tingkat kepuasan konsumen driver Gojek, model yang telah dibentuk dapat dilihat. Faktor yang menjadi perhatian selama penilaian layanan adalah kemampuan tangible dan kemampuan.

### SIMPULAN

Hasil analisis data yang dilakukan menggunakan Microsoft Excel dan metode algoritma C4.5 berdasarkan data untuk mengetahui kepuasan konsumen driver Gojek dapat dilihat sebagai berikut:

1. Jika tangible = baik dan kemampuan = cukup, maka buruk = tidak puas.
2. Jika tangible = baik dan kemampuan = cukup, maka sangat buruk = tidak puas.

### DAFTAR PUSTAKA

KASUS: PT. GO-JEK INDONESIA). *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer)*, 2(1).

<https://doi.org/10.30865/komik.v2i1.972>

Ma'ruf, A. (2017). *Analisis Kualitas Pelayanan Terhadap Kepuasan Konsumen Pada Bisnis Jasa Transportasi PT. Gojek Makassar (Studi Kasus Mahasiswa Jurusan Manajemen Angkatan 2017)*.

Magfirah, V. M. (2018). *Teknik Informatika Sekolah Tinggi Teknologi Pelita Bangsa Bekasi 2018*.

Mashlahah, S. (2017). *Jurusan Teknik Informatika Fakultas Sains Dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang 2013*.

Maulana, D., & Mustikaningtyas, V. (2022). *Penerapan Data Mining Untuk Klasifikasi Kepuasan Pelanggan Transportasi Online (Ojek Online) Menggunakan Algoritma C.4.5*.

Pasaribu, I. M. (2021). Ilmu Sosial Dan Ilmu Politik Universitas Dharmawangsa. *Jurnal Publik Reform Undhar Medan*.

Shiddiq, A., Niswatin, R. K., & Farida, I. N. (2018). *Analisa Kepuasan Konsumen Menggunakan Klasifikasi Decision Tree Di Restoran Dapur Solo (Cabang Kediri)*.

Siagian, S. P. (2020). *Sistem informasi manajemen* (6 ed.). Bumi Aksara.

Sujatmiko, A. (2020). *Pengaruh Kualitas Pelayanan Terhadap Citra Gojek Pada Mahasiswa FIKOM Universitas Islam Riau*.

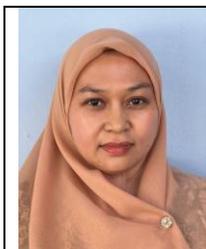
Sutabri, T. (2019). *Konsep Sistem Informasi*. Penerbit Andi.

Zahida, Q., & Sulistiyoningrum, C. E. (2018). *Analisis Penilaian Konsumen Terhadap Kinerja Layanan Transportasi Online Dalam Peningkatan Daya Saing Di Era Digital*.



Biodata:  
Penulis pertama, Frederikus Cyriaco Leu Le Uape, merupakan mahasiswa Prodi Sistem Informatika Universitas Putera Batam

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.