

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

SKRIPSI



OLEH

FREDERIKUS CYRIACO LEU LE UAPE

191510051

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
TAHUN 2023**

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

SKRIPSI

Untuk memenuhi salah satu syarat

Memperoleh gelar sarjana



OLEH

FREDERIKUS CYRIACO LEU LE UAPE

191510051

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
TAHUN 2023**

SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini saya :

Nama : Frederikus Cyriaco Leu Le Uape

NPM : 191510051

Fakultas : Teknik dan Komputer

Program Studi : Sistem Informasi

Menyatakan bahwa “**Skripsi**” yang saya buat dengan judul:

***Decision Tree* pada Penilaian Performa Layanan Driver Gojek dengan Algoritma C4.5**

Adalah hasil karya saya sendiri dan bukan “duplikasi” dari karya orang lain. Sepengetahuan, saya, didalam naskah Skripsi ini tidak terdapat karya ilmiah atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis dikutip didalam naskah ini dan disebutkan dalam sumber kutipan dan daftar Pustaka.

Apabila ternyata di dalam naskah Skripsi ini dapat dibuktikan terdapat unsur -unsur PLAGIASI, saya bersedia naskah Skripsi ini digugurkan dan gelar akademik yang saya peroleh dibatalkan, serta diproses sesuai dengan peraturan perundang -undangan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya tanpa ada paksaan dari siapapun

Batam, 3 Agustus 2023



Frederikus Cyriaco Leu Le Uape

191510051

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

SKRIPSI

Untuk memenuhi salah satu syarat

Memperoleh gelar sarjana

OLEH

FREDERIKUS CYRIACO LEU LE UAPE

191510051

**Telah disetujui oleh Pembimbing pada tanggal
seperti tertera dibawah ini**

Batam, 3 Agustus 2023



Erlin Elisa, S. Kom., M.Kom.

Pembimbing

ABSTRAK

Transportasi merupakan sarana yang penting untuk menunjang kegiatan dan operasional sehari-hari, baik itu transportasi pribadi ataupun transportasi umum. Penelitian ini akan menganalisis tingkat kepuasan pelanggan sebagai pemakai kendaraan umum, dalam hal ini penulis memanfaatkan data pelanggan yang memakai dan memanfaatkan aplikasi gojek sebagai sarana transportasi, dengan teknik pengumpulan data kuisisioner, peneliti akan mengukur tingkat kepuasan pelanggan dengan variabel *Tangible*, *Reliability*, *Responsiveness*, *Assurance*, *Empaty*. Metode analisis yang digunakan pada penelitian ini adalah datamining dengan konsep klasifikasi algoritma C4.5 yang dapat melihat hasil dari nilai *Entropy* dan *gain* tertinggi dalam bentuk pohon keputusan dengan tujuan hasil yang di peroleh dapat menjadi pedoman bagi mitra gojek untuk pelayanan dan meningkatkan pelanggan. Berdasarkan hasil klasifikasi menggunakan algoritma C4.5 menunjukkan hasil *Correctly Classified Instances* 87,21% dan *Incorrectly Classified Instances* 12,78 % dimana faktor yang menjadi penilaian kepuasan konsumen yaitu *Assurance*, *Reliability*, dan *Empaty*.

Kata kunci: *Data Mining*, *Decision Tree*, *Algoritma C4.5*, *Kepuasan Pelanggan*, *Trasnportasi online*.

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 customer data that uses and utilizes the go-jek application as a means of transportation, using a questionnaire data collection technique, the writer will measure the level of customer satisfaction with the variables Tangible, Reliability, Responsiveness, Assurance , Empathy. The analytical method used in this study is data mining with the concept of the C4.5 classification algorithm which can show the results of the highest Entrophy and gain values in the form of a decision tree with the aim that the results obtained can be a guide for gojek partners to service and increase customers. Based on the results of classicization using the C4.5 algorithm, the results show Correctly Classified Instances of 87.21% and Incorrectly Classified Instances of 12.78% where the factors that are used to evaluate customer satisfaction are Assurance, Reliability, and Empathy.

Keyword: *Data Mining, Decision Tree, Algoritma C4.5, Costumer Satisfation,*

Trasnpotation online.

KATA PENGANTAR

Puji syukur penulis mengucapkan kepada Tuhan Yang Maha Esa yang telah melimpahkan segala rahmat dan karuniaNya, sehingga penulis dapat menyelesaikan laporan tugas akhir yang merupakan salah satu persyaratan untuk menyelesaikan program studi strata satu (S1) pada Program Studi Sistem Informasi Universitas Putera Batam.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Karena itu, kritik dan saran akan senantiasa penulis terima dengan senang hati. Dengan segala keterbatasan, penulis menyadari pula bahwa skripsi ini takkan terwujud tanpa bantuan, bimbingan, dan dorongan dari berbagai pihak. Untuk itu, dengan segala kerendahan hati, penulis menyampaikan ucapan terima kasih kepada:

1. Ibu Dr. Nur Elfi Husda, S.Kom., M.SI. selaku Rektor Universitas Putera Batam.
2. Bapak Muhammat Rasid Ridho, S.Kom., M.SI. selaku Ketua Program Studi Sistem Informasi.
3. Ibu Erlin Elisa, S.Kom., M.Kom. selaku dosen pembimbing skripsi.
4. Bapak Tukino, S.Kom., M.SI. selaku dosen pembimbing akademik.
5. Dosen dan Staff Universitas Putera Batam.

Semoga Tuhan membalas kebaikan dan selalu mencurahkan hidayah serta taufik-Nya, Amin.

Batam, 3 Agustus 2023



Frederikus Cyriaco Leu Le Uape

DAFTAR ISI

HALAMAN SAMPUL	1
HALAMAN JUDUL	2
SURAT PERYATAAN ORISINALITAS	iii
HALAMAN PENGESAHAN	iv
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR	vii
DAFTAR ISI	viii
DAFTAR GAMBAR	x
DAFTAR TABEL	xi
DAFTAR RUMUS	xiii
BAB I PENDAHULUAN	
1.1. Latar Belakang.....	1
1.2. Identifikasi Masalah	4
1.3. Batasan Masalah	4
1.4. Rumusan Masalah.....	4
1.5. Tujuan Penelitian.....	5
1.6. Manfaat Penelitian.....	5
BAB II TINJAUAN PUSTAKA	
2.1. Teori Umum	6
2.2. Teori Khusus	13
2.3. Kerangka Berpikir	16
2.4. Penelitian Terdahulu.....	16
BAB III METODE PENELITIAN	
3.1. Desain Penelitian	25
3.2. Objek Penelitian	27

3.3. Populasi dan Sampel.....	27
3.3.1. Populasi	27
3.3.2. Sampel	28
3.4. Variabel Penelitian	28
3.5. Teknik Pengumpulan Data	28
3.6. Model Penelitian.....	29
BAB IV HASIL DAN PEMBAHASAN	
4.1. Hasil penelitian	31
4.2. Pembahasan	32
4.2.1 Melakukan Pra Proses Data Penelitian.....	32
4.2.2. Analisis Data	37
4.3. Analisis algoritma C4.5	45
4.4 Pengujian Software Rapid Miner	65
BAB V SIMPULAN DAN SARAN	
5.1. Simpulan.....	69
5.2. Saran	69
DAFTAR PUSTAKA	
Lampiran	73
Lampiran 1. Pendukung Penelitian.....	73
Lampiran 2. Daftar Riwayat Hidup	81
Lampiran 3. Surat Izin Penelitian	82
Lampiran 4. Surat keterangan Penerimaan Jurnal	83

DAFTAR GAMBAR

Gambar 2.1 Tahapan KDD	6
Gambar 2.2 Konsep Decision Tree	12
Gambar 2.3 Kerangka Berpikir	16
Gambar 3.1 Desain Penelitian.....	25
Gambar 3.2 Flow Chart Model Penelitian Algoritma C4.5	30
Gambar 4.1 Pohon Kepuasan Node 1	49
Gambar 4.2 Pohon Kepuasan Node 2	57
Gambar 4.3 Pohon Kepuasan Node 3	64
Gambar 4.4 Data Format Excel.....	65
Gambar 4.5 Import Data Respsitory	66
Gambar 4.6 Operator Relieble	66
Gambar 4.7 Operator Decesion Tree	67
Gambar 4.8 Pohon Keputusan Rapidminer.....	67
Gambar 4.9 Review Tree	68

DAFTAR RUMUS

2.1 Menghitung Nilai Gain	10
2.2 Menghitung Nilai Entropy	11

DAFTAR TABEL

2.1 Tabel Penelitian Terdahulu	16
4.1 Tabel Keterangan Variabel	31
4.2 Tabel Klasifikasi Tangible	38
4.3 Klasifikasi Reliability	38
4.4 Klasifikasi Responsivss.....	39
4.5 Klasifikasi Assurance.....	39
4.6 Klasifikasi Empaty	40
4.7 Format Data Akhir Penilaian Kepuasan Pelayanan Pelanggan.....	40
4.8 Data Node 1.....	48
4.9 Data Analisis Node 2	49
4.10 Node 2	56
4.11 Node 3	57
4.12 Node 3	63

BAB I

PENDAHULUAN

1.1 Latar Belakang

Saat ini, transportasi menjadi sangat penting untuk hidup karena dapat memudahkan segala aktivitas sehari-hari seperti pergi bekerja, ke sekolah, kuliah, ke pasar, jalan-jalan, dan lain-lain. Dilihat dari tingkat penggunaan kendaraan pribadi dan dinas di Indonesia, pada tahun 2020 sebanyak 51,24% pekerja menggunakan kendaraan pribadi atau dinas, sementara proporsi kendaraan umum lebih rendah, 41,93%. Proporsi ini berbeda dari tahun sebelumnya, ketika pengguna kendaraan pribadi naik menjadi 83,76% dan pengguna kendaraan umum hanya 11,81%. Namun, kendaraan umum juga menjadi pilihan bagi orang-orang yang tidak memiliki kendaraan pribadi, dan orang-orang yang sudah memiliki kendaraan juga sering menggunakan sarana transportasi umum ini.

Mengembangkan teknologi ke arah bisnis transportasi modern yang memanfaatkan kecanggihan aplikasi virtual saat ini adalah fenomena perkembangan teknologi yang paling banyak dibicarakan dan dibahas oleh banyak kalangan. Adanya sarana transportasi ini, terutama yang dapat dipesan, sangat memudahkan masyarakat saat. Dengan memiliki aplikasi dimanapun dan kapanpun, orang dapat pergi ke mana saja dengan mudah.(Zahida & Sulistiyoningrum, 2020). Semakin meningkatnya populasi dan pengembangan pemukiman di kota-kota besar menunjukkan betapa pentingnya transportasi. (Pasaribu, 2021). Transportasi online adalah salah satu metode transportasi umum yang paling umum. Sektor transportasi dan logistik online terus berkembang

sebagai akibat dari percepatan adopsi teknologi selama pandemi Covid-19 hingga saat ini. Layanan ojek online dianggap praktis dan memberikan kemudahan akses bagi penggunaannya ke mana pun dan kapan pun mereka membutuhkannya dengan dukungan teknologi terbaru. Layanan ini juga dapat meningkatkan nilai kendaraan pribadi. Studi yang dilakukan oleh *Institute for Development of Economics and Finance* (Indef) berjudul "Tren Industri Persepsi Konsumen Jasa Transportasi dan Logistik Online: Transportasi dan Logistik Online Terbukti Tangguh" juga menunjukkan hal ini.

Gojek adalah salah satu jenis transportasi online yang paling populer di Batam. PT. GOJEK adalah perusahaan yang fokus pada industri jasa transportasi, khususnya layanan ojek online. Gojek menawarkan layanan pengiriman barang, pesan antar makanan, berbelanja, dan berpergian di tengah kemacetan dengan bekerja sama dengan pengendara ojek berpengalaman. Gojek harus terus meningkatkan layanan untuk memuaskan pelanggannya dan tetap eksis di pasar transportasi online saat ini dengan banyak saingan. Untuk mengetahui sejauh mana driver Gojek telah memberikan layanan terbaik untuk pelanggannya, diperlukan penilaian kepuasan pelanggan. Menurut pengalaman penulis sebagai seorang driver, konsumen yang telah menggunakan layanan transportasi online telah diberi penilaian oleh aplikasi ini tentang pengalaman mereka menggunakan layanan tersebut. Namun, data yang terkumpul selama ini tersimpan di gudang data Gojek sendiri, sehingga driver hanya mengetahui penilaian mereka sendiri karena masing-masing driver memiliki pengalaman sendiri dalam melakukan tugasnya, sehingga untuk memahami pola penilaian konsumen ini secara keseluruhan dan untuk

meningkatkan layanan driver, perlu dilakukan analisis tingkat kepuasan konsumen, yang benar-benar merupakan faktor utama dalam menarik langganan baru.

Peneliti menggunakan data mining, yang merupakan proses yang diulang yang memerlukan interaksi manusia, untuk menemukan pola atau model baru yang dapat digeneralisasi dan digunakan untuk melakukan tindakan di masa mendatang. Dengan demikian, kepuasan pelanggan adalah kunci untuk mengetahui penilaian performa layanan driver aplikasi Gojek. (Azmi & Dahria, 2017). Perangkat tertentu yang memiliki kemampuan untuk menyediakan analisis data yang dapat dipelajari dengan teliti dan bermanfaat dapat mengidentifikasi pola-pola ini.

Peneliti menggunakan metode pohon keputusan, yang merupakan salah satu dari algoritma C4.5. Kelebihan algoritma C4.5 termasuk kemampuan untuk membuat pohon keputusan yang mudah dipahami, memiliki tingkat akurasi yang dapat diterima, dan mampu menangani data numerik dan diskret. (Adriansa dkk., 2022). Dengan menggunakan teknik data mining, penulis akan mengkategorikan penilaian performa driver Gojek untuk mengetahui tingkat kepuasan pengguna terhadap aplikasinya. Ini akan membantu langganan baru dan lama untuk terus menggunakan Gojek sebagai metode transportasi yang efisien. Berdasarkan latar belakang tersebut, penulis ingin judul tugas akhir skripsi. **“DECISION TREE PADA PENILAIAN PERFORMA LAYANAN DRIVER GOJEK DENGAN ALGORITMAC4.5”**.

1.2. Identifikasi Masalah

Dari uraian latar belakang adapun poin identifikasi masalah adalah:

1. Penilaian layanan kepuasan konsumen yang selama ini di input oleh pengguna jasa setelah memanfaatkan jasa layanan tersimpan di gudang data gojek.

1.3 Batasan Penelitian

Agar penelitian penulis tidak melenceng dari topik pembahasan, maka perlu di berikan batasan penelitian, adapun batasan penelitian sebagai berikut:

1. Penelitian ini akan melakukan analisis kepuasan pelanggan dari segi pemakai jasa layanan. Sehingga hasil analisis dapat bermanfaat bagi driver gojek untuk meningkatkan layanan mereka.
2. Analisis pengolahan data akan memanfaatkan teknik Decision Tree klasifikasi algoritma C4.5 dan melakukan pengujian hasil dengan aplikasi *opensource* Rapid.

1.4 Rumusan Masalah

Berdasarkan latar belakang diatas, maka peneliti dapat merumuskan masalah dalam penelitian ini, sebagai berikut:

1. Bagaimana data mining menganalisis data Penilaian Performa Layanan Driver Gojek dengan Algoritma C4.5 Menggunakan *Decesion Tree*?
2. Bagaimana hasil pohon keputusan yang dihasilkan setelah dilakukan perhitungan dengan nilai *gain* tertinggi yang menjadi faktor penilaian konsumen?

1.5 Tujuan Masalah

1. Untuk menganalisis data Penilaian Performa Layanan Driver Gojek dengan Algoritma C4.5 Menggunakan *Decesion Tree*.
2. Untuk mengetahui hasil pohon keputusan yang dihasilkan setelah dilakukan perhitungan dengan nilai *gain* tertinggi yang menjadi faktor penilaian konsumen.

1.6 Manfaat Penelitian

Manfaat dalam penelitian ini adalah:

1. Manfaat Teoritis

Untuk memperoleh bukti-bukti data empiris tentang analisis *Decision Tree* Pada Penilaian Performa Layanan Driver Gojek dengan Algoritma C4.5. yang akan bermanfaat untuk megembangkan ilmu pengetahuan, dan akademis mahasiswa di dalam bidang Sistem Informasi.

2. Manfaat Praktis

Penelitian ini dapat menjadi sebuah bahan informasi dan sebagai referensi tambahan yang berkaitan dengan variabel di dalam penelitian yaitu *Decision Tree* Pada Penilaian Performa Layanan Driver Gojek dengan Algoritma C4.5.

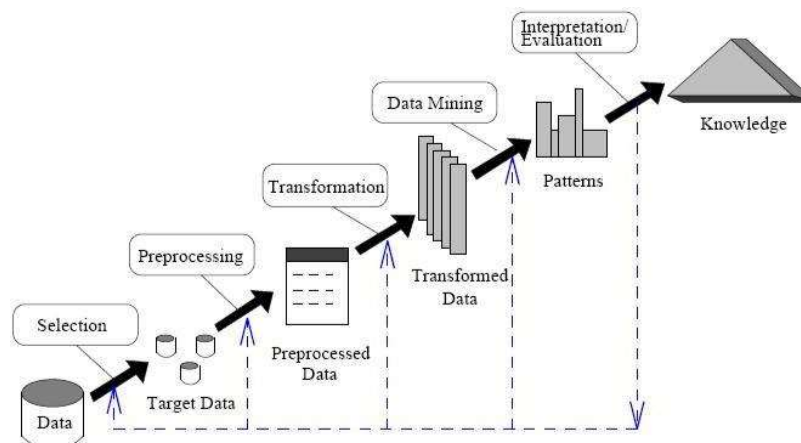
BAB II

LANDASAN TEORI

2.1 Teori Umum

2.1.1 KDD (*Knowledge Discovery in Database*)

KDD memiliki beberapa tujuan dan tahapan dalam mengumpulkan data yang besar. Tahapan Discovery of Knowledge in Database dibahas secara rinci di sini.



Gambar 2.1 Tahapan *Knowledge Discovery in Database* (KDD)

Dari gambar 2.1 diatas maka tahapan KDD hingga menjadi suatu *Knowledge* atau pengetahuan baru adalah (T. A. Q. Putri dkk., 2023).

a. *Selection*

Untuk memastikan bahwa tidak ada kesamaan dalam variable dan pengolahan data tidak mengalami kesulitan, metode pemilihan digunakan.

b. *Processing*

Processing terdapat 2 tahap, yaitu:

1. Untuk menghilangkan data yang tidak diperlukan yaitu *Data Cleaning*.

2. Data integration untuk mengidentifikasi data yang unik.

c. *Transformation*

. Proses perubahan data disebut *transformation*. Namun, data tidak akan berubah dari format yang telah ditentukan melalui proses pengolahan *datamining*.

d. *Datamining*

Setelah mendapatkan informasi dan data baru, proses utama berikutnya adalah pemrosesan, yang dilakukan sesuai dengan prosedur yang telah diterapkan untuk *datamining*.

e. *Interpretation / Evaluation*

untuk menemukan pola yang menarik data yang akan ditentukan.

f. *Knowledge*

KDD juga merupakan proses mengubah data mentah menjadi informasi yang bermanfaat untuk digunakan dalam pengambilan keputusan.

2.1.2 *Datamining*

Datamining adalah serangkaian prosedur yang digunakan untuk mengekstrak nilai tambahan dari suatu basis data, yang sebelumnya tidak diketahui secara manual. Proses ini menghasilkan informasi dengan mengekstrak dan mengidentifikasi pola yang relevan atau menarik dari data yang ada dalam basis data. Data mining terutama digunakan untuk menemukan pengetahuan dalam basis data yang besar, yang biasanya disebut *Knowledge Discovery in Database*. (KDD) (Pasaribu, 2021).

Datamining adalah teknik yang digunakan untuk mengekstrak (mengambil intisari) pengetahuan yang diperoleh dari sekumpulan data untuk mendapatkan struktur (Laila dkk., 2021). yang dapat difahami usianya dan mencakup basisdata dan manajemen data, prapemrosesan data, pertimbangan model dan inferensi, ukuran ketertarikan, pertimbangan kompleksitas, pasca pemrosesan terhadap struktur yang di temukan, visualisasi, dan peningkatan online. (Laila dkk., 2021).

2.1.3 Klasifikasi

Klasifikasi adalah metode yang bertujuan untuk mengelompokkan kumpulan data secara sistematis sesuai dengan aturan dan standar yang telah ditetapkan. Dalam data mining, klasifikasi adalah salah satu cara untuk menentukan kelas objek yang belum diketahui kelasnya. (Magfirah, 2019). Untuk tahap awal klasifikasi, akan digunakan dataset yang telah diketahui kelas objeknya untuk pelatihan dan pengujian. Proses klasifikasi didasarkan pada empat komponen utama: (Zahida & Sulistiyoningrum, 2020).

1. Kelas (*Class*)

Variabel kategori yang memberikan "label" pada objek setelah klasifikasinya dikenal sebagai variabel independen dalam model.

2. Prediktor (*Predictor*)

Variabel independen dari model ditampilkan melalui atribut atau karakteristik data yang diklasifikasikan dan didasarkan pada klasifikasi yang telah dibuat.

3. Pelatihan dataset (*Training dataset*)

Nilai-nilai dari kedua komponen sebelumnya digunakan untuk mengajarkan model untuk menemukan kelas yang tepat atau sesuai berdasarkan prediktor yang tersedia.

4. Dataset pengujian (*Testing database*)

Berisi data baru yang akan diklasifikasikan oleh klasifikator model yang telah dibangun sebelumnya. Hal ini akan memungkinkan evaluasi akurasi model prestasi (klasifikasi).

Berikut beberapa model (metode) klasifikasi yang paling populer (Zahida & Sulistiyoningrum, 2019).

- a) *Decision/classification trees.*
- b) *Bayesian classifiers/Naïve bayes classifiers.*
- c) *Neural networks.*
- d) *Statistical analysis.*
- e) *Genetic algorithms.*
- f) *Rough sets.*
- g) *K-nearest naighbor classifier.*
- h) *Rule-base methods.*
- i) *Memory based reasoning;*
- j) *Support vector machine.*

2.1.4 Algoritma C4.5

Untuk pembentukan pohon keputusan, ada banyak algoritma, seperti ID3, CART, dan C.45. Algoritma C.45 adalah pengembangan dari algoritma ID3

(Mashlahah, 2019). Algoritma C4.5, yang biasanya digunakan untuk membuat pohon keputusan, adalah sebagai berikut: (Mashlahah, 2020).

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

Nilai Gain tertinggi dari atribut yang ada adalah dasar untuk memilih atribut sebagai akar. Nilai Gain dapat dihitung dengan menggunakan rumus yang ditunjukkan dalam persamaan berikut. (Sutabri, 2019) :

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

Rumus 2.1 Menghitung Nilai *Gain*

Keterangan :

S : himpunan kasus

A : atribut

n : jumlah partisi atribut A

|S_i| : jumlah kasus pada partisi ke-i

|S| : jumlah kasus dalam S

Setelah memperoleh nilai gain, perhitungan tambahan harus dilakukan. Nilai entropy digunakan untuk menentukan seberapa informatif sebuah atribut input untuk menghasilkan atribut output. Berikut adalah rumus dasar entropi tersebut.

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

Rumus 2.2 Menghitung Nilai *Entropy*

Keterangan :

S : himpunan kasus

n : jumlah partisi S

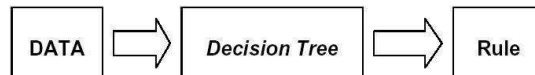
pi : proporsi dari Si terhadap S

2.1.5 *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. Dalam mengeksplorasi data, pohon keputusan sangat berguna untuk menemukan hubungan tersembunyi antara sebuah variabel target dan banyak variabel input. (Shiddiq dkk., 2019). Untuk membagi populasi yang heterogen menjadi lebih kecil, model pohon keputusan terdiri dari berbagai aturan. (homogen) dengan mempertimbangkan variabel target.

Variabel tujuan biasanya dikelompokkan dengan pasti, dan model pohon keputusan lebih mengarah pada perhitungan probabilitas dari tiap record terhadap kategori tersebut atau untuk mengklasifikasi record dengan mengelompokkannya dalam satu kelas. Sebuah pohon keputusan dapat dibangun dengan menggunakan salah satu algoritma pohon keputusan untuk memodelkan himpunan data yang

belum terklasifikasi kelasnya. (Mashlahah, 2020). Konsep dari pohon keputusan adalah mengubah data menjadi pohon keputusan dan aturan-aturan keputusan.



Gambar 2.2 Konsep *Decision Tree*

Mengubah bentuk data (tabel) menjadi model pohon, mengubah model pohon menjadi aturan, dan kemudian menyederhanakan aturan adalah proses pembuatan pohon keputusan. Langkah pertama dalam proses ini adalah menghitung nilai entropi total dari jumlah data yang dijadikan sampel. Kemudian, variabel dikelompokkan dan dihitung nilai gain untuk masing-masing atribut. Nilai gain tertinggi dari atribut tersebut akan ditentukan dengan menggunakan rumus algoritma C4.5.

2.1.6 *RapidMiner*

RapidMiner adalah perangkat lunak *opensource*. *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 mesin data mining yang dapat diintegrasikan pada produknya sendiri. *RapidMiner* berjalan di semua sistem operasi karena dibuat dengan bahasa Java.

RapidMiner menyediakan *GUI (Graphic User Interface)* untuk membuat pipeline analitis, GUI ini akan menghasilkan file XML. (Extensible Markup

Language) yang menjelaskan bagaimana pengguna ingin melakukan analisis pada data. *RapidMiner* kemudian membaca file ini, memungkinkan analisis dilakukan secara otomatis.

2.2 Teori Khusus

2.2.1 Kepuasan Pelanggan

Jika layanan penyedia tidak memenuhi harapan pelanggan, pelanggan akan merasa kecewa, tetapi jika layanan memenuhi harapan pelanggan, pelanggan akan merasa puas. Pengalaman pelanggan dengan penyedia layanan, kata-kata orang lain, dan iklan dapat menentukan harapan pelanggan..(T. A. Q. Putri dkk., 2023) . Ketidakpuasan dan kepuasan pelanggan biasanya disebabkan oleh ketidaksesuaian antara apa yang diharapkan pelanggan dan apa yang terjadi, yang berarti ada dua kemungkinan yang pasti terjadi.(Laila dkk., 2021):.

1. Kinerja yang diterima konsumen lebih baik dari yang diharapkan. Artinya konsumen puas dengan kualitas pelayanan yang diberikan oleh penyedia jasa.
2. Kinerja yang dirasakan konsumen di bawah harapan. Artinya konsumen tidak puas dengan kualitas pelayanan yang diberikan oleh penyedia jasa.

2.2.2 Kualitas Pelayanan

Seberapa jauh perbedaan antara apa yang sebenarnya terjadi dan apa yang diharapkan pelanggan tentang layanan yang mereka terima dikenal sebagai kualitas pelayanan. Kualitas layanan adalah inti dari konsep kualitas layanan, yang menunjukkan segala bentuk aktualisasi kegiatan pelayanan yang memuaskan orang-orang yang menerima pelayanan sesuai dengan daya tanggap (*respons*),

menumbuhkan adanya jaminan (*assurance*), dan menunjukkan bukti fisik (*tangible*) yang dapat dilihatnya, menurut (*empaty*), dari orang-orang yang memberikan layanan sesuai dengan keandalannya (*reliability*) menjalankan tugas layanan yang diberikan secara konsisten.(Magfirah, 2019).

a. Ketanggapan

Dalam memberikan layanan, setiap pekerja mengutamakan aspek layanan yang sangat mempengaruhi perilaku orang yang menerimanya. Oleh karena itu, mereka harus dapat melayani masyarakat dengan tingkat penyerapan, pemahaman, dan ketidaksesuaian terhadap berbagai aspek layanan yang belum mereka ketahui. Hal ini memerlukan penjelasan yang bijaksana, mendalam, membina, mengarahkan, dan membujuk untuk menyikapi semua prosedur dan mekanisme kerja yang berlaku dalam suatu organisasi, sehingga tuntutan pelayanan yang menanggapi berbagai keluhan dari bentuk pelayanan yang diberikan mendapat respons yang positif. daya tanggap pemberi pelayanan dan yang menerima pelayanan. Sayangnya, pihak yang memberikan layanan menemukan bahwa orang yang dilayani tidak memahami berbagai syarat prosedur atau mekanisme, Untuk memastikan kesan orang yang mendapat pelayanan yang benar, kesan orang yang mendapat pelayanan yang benar, dan kesan orang yang mendapat pelayanan memahami atau tanggap terhadap keinginan orang yang dilayani, perlu diberikan pemahaman dan pemahaman yang jelas dan bijaksana.(Magfirah, 2019).

b. Jaminan dan Kepastian

Setiap jenis layanan memerlukan kepastian atas layanan yang diberikan. Kepastian ini sangat ditentukan oleh jaminan dari pegawai yang memberikan

layanan, sehingga orang yang menerima layanan semakin puas dan yakin bahwa segala urusan mereka akan tuntas dan selesai dengan kecepatan, ketepatan, kemudahan, kelancaran, dan kualitas yang diinginkan..

c. Perwujud

Dengan kata lain, kemampuan untuk menciptakan ekstensi yang melibatkan elemen eksternal. Penampilan dan kemampuan sarana dan prasarana fisik yang dapat diandalkan keadaan lingkungan sekitar adalah bukti nyata dari layanan yang diberikan oleh driver transportasi online. Contoh peralatan dan peralatan (teknologi) yang digunakan serta penampilan driver merupakan contoh sarana dan prasarana fisik yang dapat diandalkan.

d. Keandalan

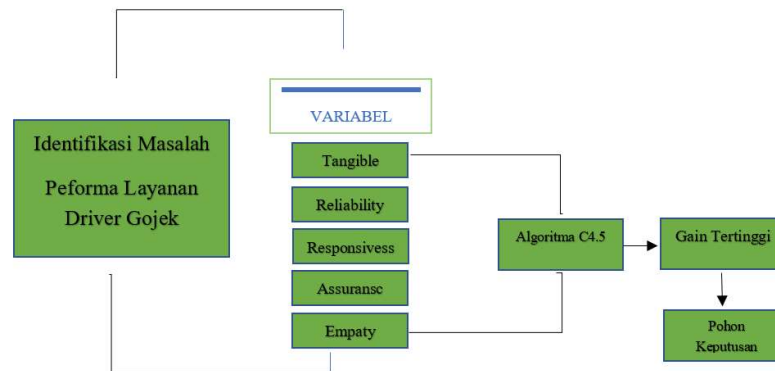
Dengan kata lain, kemampuan driver untuk memberikan layanan yang dijanjikan dengan tepat dan aman. Kinerja harus memenuhi harapan pelanggan, yang berarti ketepatan waktu, layanan yang sama untuk semua pelanggan, ramah, dan akurat..

e. Empati

Dengan kata lain, memberikan perhatian yang tulus dan bersifat individual atau pribadi kepada pelanggan dengan berusaha memahami keinginan pelanggan. Driver transportasi online (Go-Jek) diharapkan mengenal dan memahami pelanggan, memahami kebutuhan khusus pelanggan, dan membuat perjalanan menjadi nyaman.

2.3 Kerangka Berpikir

Peneliti membuat gambaran singkat dengan kerangka pemikiran sebagai berikut:



Gambar 2.3 Kerangka Berpikir

Sumber: (penulis 2023)

2.4 Penelitian Terdahulu

Berikut ini beberapa penelitian terdahulu yang menjadi acuan peneliti dalam melakukan penelitian ini

Tabel 2.1 Tabel Penelitian Terdahulu

No	Peneliti	Judul Peneliti	Hasil
1	(Adriansa dkk., 2022)	Analisis Kepuasan Pelanggan Menggunakan Algoritma C4.5	Analisis sistem yang dibuat menggunakan bahasa pemrograman visual basic 2010 dan metode yang digunakan yaitu Algoritma C4.5. Teknik pengambilan data yang didapat menggunakan metodologi observasi, wawancara, dan studi pustaka. Tahapan dalam sistem kepuasan dengan metode algoritma C4.5 yaitu pengambilan data pelanggan dan data atribut dari kuisisioner, perhitungan nilai entropy dan gain, serta hasil kepuasan. Hasil yang diperoleh dari kuisisioner yang

			dibagikan kepada 20 pelanggan terdapat 15 pelanggan dengan hasil puas dan 5 pelanggan dengan hasil tidak puas. Sehingga dari sistem yang dibuat dapat membantu pihak pangkalan untuk mengetahui tingkat kepuasan para pelanggan menggunakan sistem dengan metode algoritma C4.5.
2	(Maulana & Mustikaningtyas, 2022)	Penerapan Data Mining Untuk Klasifikasi Kepuasan Pelanggan Transportasi Online (Go-Jek) Menggunakan Algoritma C.4.5	Di Indonesia banyak terdapat sepeda motor yang juga berfungsi sebagai kendaraan umum yaitu mengangkut orang/barang. Saat ini sudah banyak penyedia jasa transportasi online (Go-Jek) 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.
3	(Azwanti & Elisa, 2020)	Analisa Kepuasan Konsumen Menggunakan Algoritma C4.5	Algoritma C4.5 dianggap sebagai salah satu algoritma Data Mining yang mampu memberikan hasil yang baik karena menghasilkan aturan pohon keputusan. Hasil penelitian menunjukkan bahwa variabel Rasa mendapat Gain tertinggi pertama

			dengan nilai 0.3659. perhitungan selanjutnya didapat variabel Kebersihan dengan nilai Gain 1. Rule yang dihasilkan sebanyak 4, 2 rule menghasilkan keputusan Puas dengan kondisi jika Rasa Enak, maka Konsumen Puas dan jika Rasa Lumayan dan Kebersihan Bersih maka Konsumen Puas. Tujuan dari analisis ini supaya usaha Bakso Akbar dapat mempertahankan pelanggannya dan memperbaiki kekurangannya berdasarkan dari penilaian konsumen.
4	(Simajuntak & Erlin, 2019)	Data Mining Untuk Menentukan Pemilihan Celular Card Di Kota Batam	Tingkat penggunaan kartu oleh pelanggan seluler sangat didominasi oleh pelanggan kartu Prabayar. Pelanggan kartu Prabayar bisa dengan mudahnya berpindah ke operator lain dengan registrasi ulang kartu. Kartu Perdana Prabayar saat ini dengan mudah didapatkan dimanapun dengan harga yang sangat terjangkau dan rata-rata murah. Penawaran dari semua operator seluler rata-rata sama baik kualitas, harga atau layanannya. Tujuan penelitian ini adalah untuk melakukan pemilihan celuler card di Kota Batam. Penyelesaian masalah ini menggunakan Data Mining dengan metode Asosiasi dengan algoritma Apriori. Dengan implementasi software menggunakan software Tanagra 4.1.
5	(Marlina & Bakri, t.t.)	Penerapan Data Mining Untuk Memprediksi Transaksi Nasabah Dengan Algoritma C4.5	Penelitian ini di khususkan untuk mengklasifikasi dalam memprediksi transaksi nasabah yang akan nunggak atau lancar dalam pembayaran angsuran pada koperasi BMT Sahabat Rawajitu Selatan. Adapun masalah yang di hadapi pada penelitian ini adalah koperasi BMT Sahabat Rawajitu Selatan dalam melakukan analisis-analisis setiap transaksi

			<p>nasabahnya sangat lah kurang. Hal ini dapat dilihat dari adanya nasabah-nasabah yang nunggak pembayaran angsuran bahkan sampai ada nasabah yang sulit dalam melakukan pembayaran angsurannya. Metode yang digunakan dalam penelitian ini adalah klasifikasi data adalah C4.5 dengan menggunakan data primer yang didapat dari koperasi BMT Sahabat Rawajitu Selatan untuk proses klasifikasi.</p>
6	(Umam dkk., 2020)	Penerapan Algoritma C4.5 Untuk Prediksi Loyalitas Nasabah PT Erdika Elit Jakarta	<p>Algoritma C4.5 merupakan kelompok algoritma decision tree. Algoritma ini mempunyai input berupa training samples dan samples. Sedangkan samples merupakan field-field data yang nantinya akan kita gunakan sebagai parameter dalam melakukan klasifikasi data. Dari variabel frekuensi transaksi perusahaan dapat melihat nasabah mana yang loyal terhadap perusahaan berdasarkan data historis transaksi nasabah, namun masih ada beberapa variabel yang membuat nasabah loyal terhadap perusahaan. Variabel – variabel tersebut ialah usia, gender nasabah, gender sales perusahaan, latar belakang pendidikan, frekuensi transaksi nasabah. Perusahaan mengetahui bagaimana memprediksi nasabah yang akan loyal terhadap perusahaan berdasarkan pengalaman dari beberapa variabel diatas, namun perusahaan tidak mengetahui variabel yang paling berpengaruh dalam penilaian nasabah yang loyal karena dari beberapa variabel diatas tidak saling terhubung dan belum pasti kalau satu variabel dapat membuat keputusan apakah nasabah tersebut loyal. Berdasarkan pohon keputusan yang telah</p>

			dibuat atribut yang paling berpengaruh terhadap loyalitas nasabah adalah latar belakang pendidikan karena memiliki nilai gain yang paling tinggi yaitu 1.545292721 dan sebagai akar dari pohon keputusan sedangkan gender nasabah tidak terlalu berpengaruh terhadap loyalitas nasabah karena selalu berada pada node terakhir dengan nilai gain yaitu 0,623919119.
7	(Febriarni & Astuti, 2019)	Penerapan Algoritma C4.5 untuk Prediksi Kepuasan Penumpang Bus Rapid Transit (BRT) Trans Semarang	Dalam penelitian ini variabel pengukuran kepuasan penumpang yang digunakan adalah harga, fasilitas, dan pelayanan. Dari penelitian ini atribut yang digunakan menghasilkan hubungan sebab-akibat dalam mengklasifikasikan penumpang puas dan tidak puas. Penelitian ini diharapkan dapat membantu Pemerintah dalam meningkatkan kepuasan penumpang agar masyarakat beralih menggunakan BRT Trans Semarang dan dapat memberikan pengaruh terjadinya peningkatan jumlah penumpang. Berdasarkan hasil klasifikasi dengan menggunakan algoritma C4.5 menunjukkan adanya peningkatan di tiap pengujian yang telah dilakukan dengan hasil akurasi akhir yang cukup baik sebesar 95% yang menunjukkan bahwa algoritma C4.5 cocok digunakan untuk mengukur tingkat kepuasan penumpang BRT Trans Semarang.
8	(Santoso & Sekardiana, 2019)	Penerapan Algoritma C4.5 Untuk Penentuan Kelayakan Pemberian Kredit (Studi Kasus: Koperasi-Koperasi)	Penelitian ini bertujuan membentuk model klasifikasi pohon keputusan untuk menentukan kelayakan kredit nasabah. Pada penelitian ini penerapan Algoritma C4.5, berdasarkan Himpunan dan Atribut yang digunakan pada penelitian ini yaitu, Jumlah Penghasilan dibagi menjadi 2 kategori >5 Juta dan 3-5Juta, Jumlah Saldo dibagi menjadi tiga

		Warga Komplek Gandaria)	yaitu >3 Juta, 1-3 Juta dan < 1Juta, Jumlah Pinjaman dibagi menjadi tiga yaitu 1-4 Bulan, 5-8 bulan, dan 9-12 Bulan dan Keperluan dengan attribute Modal Usaha, membeli barang dan lain-lain. Pada Penelitian ini menentukan node akar yang sesuai, hasil klasifikasi menggunakan Algoritma C4.5 menunjukkan bahwa diperoleh akurasi 97,5%, berdasarkan hasil yang diperoleh menunjukkan bahwa algoritma c4.5 cocok digunakan untuk menentukan kelayakan pemberian kredit nasabah pada KOPERIA.
9	(Riandari & Simangunsong, 2019)	Penerapan Algoritma C4.5 Untuk Mengukur Tingkat Kepuasan Mahasiswa	Mengukur kepuasan mahasiswa sangatlah penting untuk diperhatikan mengingat tingginya tingkat persaingan dalam dunia pendidikan seiring berkembangnya pengetahuan dan teknologi. Dalam mengukur kepuasan mahasiswa sangatlah penting diperhatikan apakah pelayanan yang diharapkan oleh mahasiswa sesuai dengan yang diterima. Mengukur kepuasan mahasiswa akan sangat membantu suatu Perguruan Tinggi dalam peningkatan kualitas pelayanan yang nantinya juga akan berdampak pada peningkatan jumlah mahasiswa. Sistem yang sangat dibutuhkan adalah Penerapan Data Mining Dalam Mengukur Kepuasan Mahasiswa. Dimana mahasiswa akan menjadi objek yang memberikan penilaian/opini terhadap variable-variabel yang memiliki karakteristik yaitu Tangible, Reability, Responsivnes, Assurance, Empathy. Sistem ini dibangun dengan algoritma C4.5 dan Pengujian system dengan batuan aplikasi RapidMiner 7.5

10	(Yulianti dkk., 2020)	Optimasi Akurasi Algoritma C4.5 Berbasis Particle Swarm Optimization dengan Teknik Bagging pada Prediksi Penyakit Ginjal Kronis	Penerapan model algoritma C4.5 individual dan algoritma C4.5 berbasis PSO dengan teknik bagging dilakukan guna mengetahui model mana yang memberikan hasil terbaik dalam mendeteksi penyakit ginjal kronis. Pemilihan kedua model tersebut dipertimbangkan karena algoritma C4.5 merupakan salah satu algoritma terbaik data mining, namun cenderung memiliki kelemahan pada data yang overlap, kelas dan atribut yang banyak. Maka dari itu, teknik optimasi Particle Swarm Optimization (PSO) dan teknik bagging juga dipilih sebagai alternatif dalam mengatasi kelemahan yang ada dalam algoritma C4.5. Dari hasil penelitian diperoleh bahwa model algoritma C4.5 berbasis PSO dengan teknik bagging mampu menyeleksi atribut sehingga dapat meningkatkan nilai akurasi yang lebih baik dengan hasil sebesar 99,70% dibandingkan dengan model individual algoritma C4.5 yang menghasilkan nilai akurasi sebesar 91,72% saja.
11	(Febriyani dkk., 2021)	Index Kepuasan Pelanggan Informa dengan Menggunakan Algoritma C.45	Penelitian ini menggunakan 6 (enam) Kriteria dan 21 (Dua Puluh Satu) Sub Kriteria dengan rincian sebagai berikut Kriteria yaitu Staff, Produk, Kenyamanan Toko, Kemanan, Pengiriman dan Instalasi. Tujuan penelitian ini sebagai bahan evaluasi pihak manajemen dalam meningkatkan penelitian. Tahapan penelitian mengacu pada Knowledge Discovery Data yaitu Data, Data Selection, Data Transformation, Model data mining, Intrepretation. Hasil Akurasi akurasi yang didapat yaitu sebesar 84,16 % dengan rincian sebagai berikut

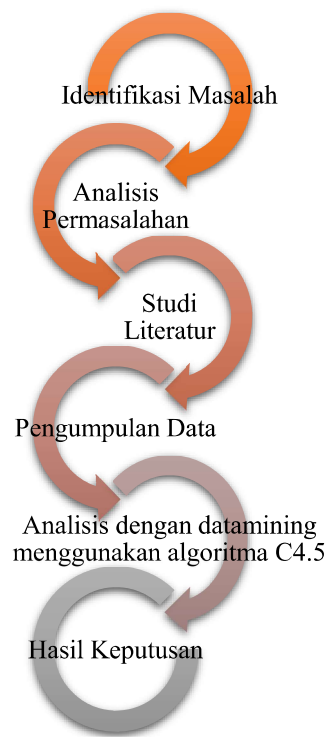
			<p>Hasil Prediksi Puas dan ternyata True Puas Sebesar 80 Data, Hasil Prediksi Puas dan ternyata True Sangat Puas Sebesar 15 Data, Hasil Prediksi Puas dan ternyata True Cukup Puas Sebesar 0 Data, Hasil Prediksi Sangat Puas dan ternyata True Puas Sebesar 0 Data, Hasil Prediksi Sangat Puas dan ternyata True Sangat Puas Sebesar 0 Data, Hasil Prediksi Sangat Puas dan ternyata True Cukup Puas Sebesar 0 Data.</p>
12	(S. U. Putri dkk., 2021)	Implementasi Data Mining Untuk Prediksi Penyakit Diabetes Dengan Algoritma C4.5	<p>Pada penelitian ini bertujuan untuk membuat model prediksi menggunakan Data Mining Algoritma C4.5 yang menghasilkan sebuah pohon keputusan serta pengujian yang dilakukan dengan menggunakan Rapidminer agar pencegahan terhadap penyakit diabetes dapat dilakukan segera mungkin. Dalam penelitian ini terdapat beberapa atribut klasifikasi yaitu berat badan, usia, tekanan darah, denyut nadi dan kadar gula darah. Hasil dari penelitian ini akan dijadikan sebagai acuan untuk dapat melihat apakah seseorang beresiko terkena penyakit diabetes atau tidak berdasarkan atribut yang telah ditetapkan.</p>
13	(Rajeswari dkk, 2019)	A Gear fault identification using wavelet transform, rough set based GA, ANN and C4.5 algorithm	<p>Early fault detection methodology in gear box diagnosis has been proposed to find the status of the gear based on vibration signals obtained from the experimental test rig. Signal processing categorized to time-frequency domain such as continues wavelet transform is used in the proposed work for statistical feature extraction. Feature selection method is used for selecting the extensive useful features among the extracted features to reduce the processing time</p>

14	(Huan-Bin Wang & Yang Jun Gao, 2020)	Research on C4.5 algorithm improvement strategy based on MapReduce	By analyzing the shortcomings of C4.5 decision tree algorithm, this paper combines MapReduce parallel model in Hadoop platform with C4.5 decision tree algorithm, so that C4.5 algorithm can be executed in parallel, so as to improve the efficiency of the algorithm. The improved C4.5 algorithm implementation process is given and the process is analyzed. Finally, the C4.5 algorithm improved experimental environment configuration is explained
15	(Ching-kan lo dkk, 2019)	Smart Dynamic Resource Allocation Model for Patient-Driven Mobile Medical Information System Using C4.5 Algorithm	A dynamic resource allocation table is critical for optimizing the performance to the mobile system, including the doctors, nurses, or other relevant health workers. We have designed a smart dynamic resource allocation model by using the C4.5 algorithm and cumulative distribution for optimizing the weight of resource allocated for the five major attributes in a cooperation communications system.

BAB III METODE PENELITIAN

3.1 Desain Penelitian

Dalam pelaksanaan penelitian *Decision Tree* Pada Penilaian Performa Layanan Driver Gojek Dengan Algoritma C4.5 digunakan rencana yang berisi langkah-langkah untuk memberikan arah dalam membahas permasalahan penelitian. Gambaran mengenai desain dari penelitian yang dilakukan, sebagai berikut.



Gambar 3.1. Desain Penelitian

Sumber: (Penulis 2023)

Pada gambar mengenai desain penelitian di jelaskan sebagai berikut:

1. Identifikasi Masalah

Pada tahap ini diuraikan mengenai permasalahan yang terdapat pada penelitian, yaitu bagaimana data mining menganalisis data Gojek dengan Algoritma C4.5 menggunakan *decesion tree* dan bagaimana hasil pohon keputusan yang dihasilkan setelah dilakukan perhitungan dengan nilai gain tertinggi yang menjadi faktor penilaian konsumen.

2. Analisis Permasalahan

Tahap ini merupakan kegiatan lanjutan setelah menjelaskan masalah penelitian. Setelah mendapatkan permasalahan, maka ditetapkan keputusan untuk menganalisis data penilaian performa layanan driver Gojek sehingga informasi yang dihasilkan dapat menjadi pedoman dalam membuat kebijakan bagi Driver Gojek sehingga dapat lebih meningkatkan penilaian konsumen.

3. Studi Literatur

Untuk mendukung pelaksanaan kegiatan penelitian, maka dilakukan proses pencarian informasi atau studi literatur dari beberapa referensi seperti jurnal penelitian yang membahas mengenai analisa penilaian performa layanan serta penelitian terdahulu yang berkaitan dengan penelitian yang akan diteliti dan juga buku-buku yang berkaitan dengan algoritma C4.5 dan beberapa indikator yang dapat memberikan pengaruh kepada penilaian performa layanan driver Gojek.

4. Pengumpulan Data

Proses dalam pengumpulan data dilakukan dengan memberikan kuesioner kepada masyarakat yang pernah menggunakan layanan Gojek.

5. Analisis Dengan Data mining Menggunakan Algoritma C4.5

Data dianalisis dengan data mining menggunakan algoritma C4.5 dan Selanjutnya dilakukan perhitungan nilai gain tertinggi dan hasilnya menjadi acuan dalam proses pembuatan struktur *decision tree*.

6. Hasil

Setelah dilakukan analisis maka penulis akan menguji hasil dengan software rapidminer untuk melakukan validasi hasil perhitungan manual sehingga disini nampak pohon keputusan.

3.2 Objek Penelitian

Pada sub bab objek penelitian, penulis menjelaskan keterangan lokasi penelitian dan penjelasan jadwal selama kegiatan penelitian, sebagai berikut.

3.3 Lokasi Penelitian

Lokasi dalam penelitian ini adalah kota Batam Provinsi Kepulauan Riau

3.3.1 Populasi dan Sampel

Dalam sub bab ini penulis akan menjelaskan mengenai keterangan yang dimiliki oleh populasi dan sampel yang dijadikan bahan dalam pelaksanaan kegiatan penelitian sebagai berikut:

3.3.1 Populasi

Populasi penelitian ini adalah seluruh masyarakat yang pernah melakukan pemesanan dan pemanfaatan pelayanan aplikasi Gojek. Dalam hal populasi selama melakukan penyebaran kuisisioner penulis hanya memperoleh 133 responden.

3.3.2 Sampel

Karena populasi data yang diperoleh dari hasil kuisisioner untuk itu penulis akan memanfaatkan metode sampling jenuh dimana seluruh populasi digunakan sebagai sampel, sehingga data yang akan di analisis sebanyak 133 data.

3.4 Variabel Penelitian

Beberapa indikator yang digunakan dalam penelitian khususnya yang berkaitan dengan variabel dari kualitas layanan, yaitu:

1. *Tangible*
2. *Reliability*
3. *Responsiveness*
4. *Assurance*
5. *Empathy*

3.5. Teknik Pengumpulan Data

Proses pengumpulan data yang menjadi kebutuhan bahan penelitian dilakukan agar penulis dapat mencapai tujuan penelitian, yaitu:

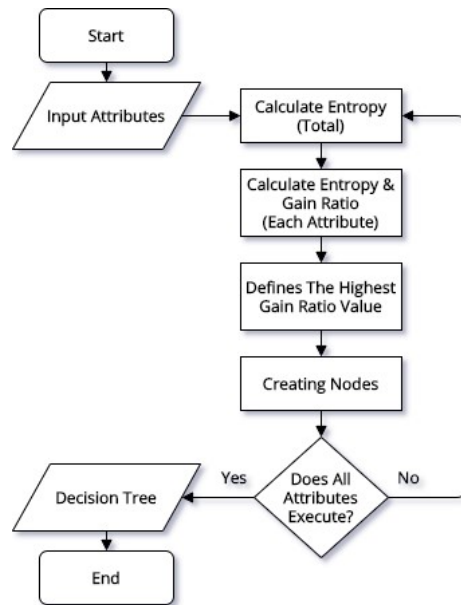
1. Studi Literatur
Mempelajari bahan dari jurnal penelitian yang sudah dilakukan peneliti lainnya namun tetap berkaitan dengan penelitian yang dilakukan penulis.
2. Observasi
Mengamati secara langsung konsumen yang menggunakan jasa aplikasi Gojek secara langsung di Batam Center.
3. Kuesioner

Proses kegiatan dengan memberikan kuesioner berisi pertanyaan menggunakan indikator dari variabel kualitas layanan kepada para konsumen pengguna layanan aplikasi Gojek.

3.6 Model Penelitian

Penulis melakukan penelitian, khususnya kegiatan analisis terhadap data menggunakan proses *Knowledge Discovery In Database* (KDD) dengan data mining, *decision tree* dan algoritma C4.5. Tahapan awal dilakukan klasifikasi sesuai dengan kategori atau kriteria data, selanjutnya dilakukan perhitungan algoritma dari C4.5 untuk mendapatkan nilai *entropy* dan *gain* tertinggi pada saat membuat struktur dari *decision tree* dilakukan secara berulang hingga simpul pada decision tidak memiliki turunan lagi sampai didapatkan pola berupa aturan dasar penilaian mengenai performa layanan driver gojek. Pada penelitian, digunakan KDD (*Knowledge Discovery In Database*), *decision tree* dan algoritma C4.5.

Untuk melihat lebih jelas model penelitian ini dapat di gambarkan melalui *flowchart* berikut.



Gambar 3.2 *Flowchart* Model Penelitian Algoritma C4.5