E-ISSN : 2988-1986
https://ejournal.cahayailmubangsa.institute/index.php/kohesi

Kohesi: Jurnal Multidisiplin Saintek
Volume 10 No 7 Tahun 2025

KLASIFIKASI EFEKTIVITAS FUNGSIONAL KEPUASAN PENGGUNA MRT GEN Z DENGAN METODE DATA MINING ALGOROTMA C4.5

Zaky Gymnastiar[1*], Muhammad Fikry Fadhillah[2], Inayah Sari[3], Aulia Rossa Nasution[4]
[1,2,3,4] Program Studi Sistem Informasi, Fakultas Teknik Informatika, Universitas Bina Sarana Informatika, Jl. Raya Kaliabang Tengah No8, Bekasi, Jawa Barat, Indonesia
E-mail: zaky.gymnastiar0418@gmail.com[1], fikram1714@gmail.com[2], sariinayah61@gmail.com[3], liaanst16@gmail.com[4]

ABSTRACT

MRT or Mass Rapid Transit is a form of public transportation designed to transport large numbers of passengers at high speeds. With the increasing use of modern public transportation such as MRT among Gen Z, an evaluation of the operational effectiveness of this service was conducted. Functional effectiveness relates to how well the MRT can meet the needs of users. This study aims to identify the level of satisfaction of Gen Z MRT users using a data mining method that applies the C4.5 algorithm. The factors analyzed include comfort, safety, and punctuality. Data was collected through a questionnaire distributed to Gen Z, using 100 respondents as a research sample and processed using RapidMiner. The test results show that the classification model developed was able to achieve an accuracy rate of 97%, indicating the effectiveness of the C4.5 algorithm in grouping user satisfaction levels. Based on the results of the study, it can be concluded that the data mining method using the C4.5 algorithm successfully classified the level of satisfaction of Gen Z MRT users effectively. The resulting decision tree structure shows that punctuality is the main factor in determining MRT user satisfaction.

Keywords: MRT, Gen Z, Data Mining, C4.5 Algorithm, Classification

ABSTRAK

MRT atau Mass Rapid Transit merupakan transportasi umum yang dirancang untuk pengangkut penumpang dalam jumlah besar dengan kecepatan tinggi. Dalam peningkatan pemanfaatan transportasi umum modern seperti MRT di kalangan Gen Z, maka dilakukan evaluasi terhadap efektivitas operasional layanan ini. Efektivitas fungsional berkaitan dengan seberapa baik MRT dapat memenuhi kebutuhan pengguna. Penelitian ini bertujuan untuk mengidentifikasi tingkat kepuasan pengguna MRT pada Gen Z menggunakan metode data mining yang mengaplikasikan algoritma C4.5. Faktor-faktor yang dianalisis meliputi kenyamanan, keamanan, dan ketepatan waktu. Data dihimpun melalui penyebaran kuesioner kepada kalangan Gen Z, serta meggunakan 100 responden yang dijadikan sebagai sample penelitian dan diolah dengan menggunakan alat RapidMiner. Hasil dari pengujian menunjukkan bahwa model klasifikasi yang dikembangkan mampu

Article History
Received: Desember 2025
Reviewed:Desember 2025
Published: Desember 2025

Plagiarism Checker No 235
Prefix DOI : 10.8734/Kohesi.v1i2.365
Copyright : Author
Publish by : Kohesi

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

mencapai tingkat akurasi sebesar 97%, yang mengindikasikan efektivitas algoritma C4.5 dalam mengelompokkan tingkat kepuasan pengguna. Berdasarkan hasil penelitian dapat disimpulkan metode data mining dengan algoritma C4.5 berhasil mengklasifikasikan tingkat kepuasan pengguna MRT dari kelompok Gen Z secara efektif. Struktur decision tree yang dihasilkan menunjukkan bahwa ketepatan waktu adalah faktor utama dalam menentukan kepuasan pengguna MRT.

Kata Kunci: ***MRT, Generasi Z, Penambangan Data, Algoritma C4.5, Klasifikasi***

1. PENDAHULUAN

Transportasi umum di Jakarta merupakan komponen penting yang mendukung berbagai aktivitas sosial dan ekonomi masyarakat perkotaan. Dengan meningkatnya jumlah perjalanan yang dilakukan masyarakat setiap hari, transportasi umum berperan signifikan dalam mewujudkan sistem mobilitas yang lebih efisien. Sistem transportasi yang efisien berkontribusi terhadap peningkatan konektivitas antar kawasan pusat bisnis, area permukiman, dan tempat wisata di Jakarta (Putri, 2022). Oleh karena itu pemerintah telah berupaya mengembangkan berbagai moda transportasi umum, salah satu moda transportasi umum yang telah dikembangkan pemerintah untuk meningkatkan efisiensi mobilitas tersebut adalah *Mass Rapid Transit* (MRT).

Mass Rapid Transit (MRT) adalah proyek penting yang dimulai oleh pemerintah untuk menciptakan sistem transportasi modern di Jakarta, dengan tujuan mengurangi kemacetan dan meningkatkan kenyamanan dalam bepergian. Sejak beroperasi pada tahun 2019, MRT Jakarta menjadi contoh perubahan dalam sistem transportasi dengan menekankan kecepatan, ketepatan waktu, kenyamanan, serta ramah lingkungan. Menurut data BPS DKI Jakarta, jumlah pengguna MRT pada bulan Januari 2025 mencapai 3.534.665 orang (Statistik, 2025), yang menunjukkan tingkat kepercayaan masyarakat terhadap layanan yang diberikan. Namun, masih perlu dilakukan penelitian lebih lanjut untuk menganalisis sejauh mana MRT benar-benar berhasil meningkatkan efisiensi perjalanan dan kepuasan pengguna.

Dilakukannya penelitian karena penting untuk mengetahui tingkat kepuasan pengguna MRT terhadap pelayanan yang telah diberikan, yang diproses menggunakan metode tertentu dan ditulis dalam karya ilmiah. Penilaian layanan MRT yang berfokus pada aspek kenyamanan, keamanan, dan ketepatan waktu. Metode yang digunakan untuk memproses data adalah algoritma C4.5 dengan menggunakan tools *RapidMiner*, analisis yang dihasilkan mengenai apakah pengguna merasa puas atau tidak dengan kualitas layanan MRT.

Kepuasan pelanggan adalah tujuan utama dari keberhasilan sebuah perusahaan, oleh karena itu perusahaan perlu memahami dan memenuhi keinginan serta harapan masyarakat. Kepuasan pelanggan adalah perasaan yang dialami seseorang, bisa berupa senang atau kecewa, setelah membandingkan apa yang telah diperoleh dengan apa yang diharapkan. Keberhasilan sebuah perusahaan, termasuk dalam sektor transportasi umum, bergantung pada kepuasan pelanggan dan kinerja karyawan. Kedua faktor ini menjadi dasar utama dalam mencapai kesuksesan perusahaan (Purnandika & Septiana, 2023).

Kualitas pelayanan dalam transportasi umum merupakan aspek penting yang berperan dalam membentuk tingkat kepuasan serta kepercayaan pengguna terhadap sistem transportasi publik. Pelayanan yang baik mencerminkan kemampuan penyelenggara transportasi dalam memenuhi kebutuhan dan ekspektasi pengguna melalui berbagai dimensi, seperti ketepatan jadwal, kenyamanan fasilitas, dan keamanan perjalanan. Peningkatan mutu pelayanan tidak hanya menciptakan persepsi positif masyarakat terhadap transportasi umum, tetapi juga berpengaruh terhadap peningkatan efisiensi mobilitas serta pengurangan ketergantungan masyarakat pada kendaraan pribadi.

Adapun penelitian yang serupa, penelitian ini berjudul “Perbandingan Hasil Penerapan Algoritma Klasifikasi dan NLP Terhadap Data Kepuasan Pengguna MRT Jakarta”. Penelitian dilakukan dengan menganalisis 525 unggahan dari media sosial X yang sudah diberi label positif, negatif, dan netral. Tujuannya adalah untuk memahami persepsi masyarakat terhadap MRT Jakarta serta mengetahui algoritma klasifikasi yang paling efektif. Data dibagi dalam rasio 80:20, dan hasil menunjukkan bahwa *Random Forest* memiliki akurasi tertinggi, yaitu 81%, diikuti oleh *Support Vector Machine* (78%) dan *Logistic Regression* (74%) (Pribadi & Ernawati, 2024). Kelebihan penelitian ini adalah pengumpulan data yang dilakukan secara manual dan terstruktur, serta proses pra-pemrosesan yang lengkap seperti pembersihan dan tokenisasi. Penelitian juga membandingkan tiga algoritma berbeda dan melakukan penyesuaian parameter. Namun, ada beberapa kelemahan, seperti potensi bias karena data berasal dari pengguna media sosial, ketidakseimbangan kelas, serta kurangnya informasi terperinci mengenai evaluasi kinerja model.

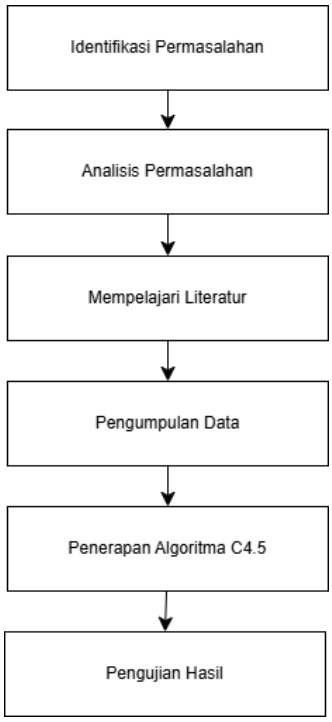
Pada penelitian sebelumnya yang menggunakan Tiga algoritma yang diuji *Random Forest*, *Support Vector Machine*, dan *Logistic Regression*. Dengan demikian penggunaan algoritma C4.5 menjadi pembeda pada penelitian ini yang bertujuan untuk mengidentifikasi tiga faktor yang paling berpengaruh terhadap kepuasan pengguna MRT Jakarta yang berfokus pada faktor kenyamanan, keamanan, dan ketepatan waktu, serta menghasilkan model *decision tree* yang dapat memberikan pemahaman komprehensif tentang pola hubungan antara variabel-variabel

tersebut, penelitian ini juga memanfaatkan data terbaru tahun 2025 mengenai tingkat kepuasan pengguna MRT Jakarta agar hasil yang diperoleh lebih relevan.

2. METODE PENELITIAN

2.1 Tahapan Penelitian

Tahapan dalam penelitian ini digambarkan melalui bagan pada gambar berikut:



Gambar 1. Tahapan Penelitian

Penjelasan mengenai gambar di atas adalah sebagai berikut:

1. Identifikasi Permasalahan

   Sebelum penelitian dilakukan, penting untuk mengidentifikasi permasalahan penelitian secara jelas. Permasalahan dalam penelitian ini adalah bagaimana mengklasifikasi tingkat kepuasan pengguna terhadap efektivitas fungsional MRT, khususnya pada Gen Z.

2. Analisis Permasalahan

   Untuk memperoleh hasil analisis yang lebih detail. Setelah permasalahan diidentifikasi, penting untuk menentukan batasan atau ruang lingkup dari penelitian.

3. Mempelajari Literatur

   Jurnal dan buku yang mengangkat isu kepuasan pelanggan dengan metode *data mining* telah dikumpulkan sebagai acuan untuk penelitian.

4. Pengumpulan Data

   Data tersebut diperoleh dari kuesioner yang dibuat menggunakan media *Google Form* dengan skala *guttman* lalu disebarkan kepada pengguna MRT, khususnya dari kalangan Gen Z. Dalam kuesioner tersebut mencakup berbagai faktor yangmempengaruhi

Tingkat kepuasan pengguna, seperti kenyamanan, keamanan, dan ketepatan waktu dalam penggunaan layanan MRT.

5. Penerapan Algoritma C4.5

   Tahapan ini mencakup proses penerapan algoritma C4.5 untuk mengolah data hingga menghasilkan pohon keputusan akhir yang menggambarkan hasil klasifikasi.

6. Pengujian Hasil

   Untuk menilai seberapa akurat hasil analisis, penerapan algoritma C4.5 akan diuji coba dengan menggunakan *software RapidMiner Studio*.

2.2 *Data Mining*

Istilah *data mining* merujuk pada aktivitas atau proses untuk menemukan pola, keterkaitan, atau wawasan baru dari kumpulan data besar yang sebelumnya tidak diketahui, dengan maksud untuk membantu dalam pengambilan keputusan yang lebih baik. Secara lebih luas, *data mining* merupakan proses untuk menemukan informasi penting dari sejumlah data. Informasi ini diperoleh melalui proses yang kompleks seperti pemanfaatan kecerdasan buatan, metode statistik, matematika, *machine learning*, dan lain-lain. Metode-metode yang rumit tersebut nantinya berfungsi untuk mengenali dan mengambil informasi yang berguna dari basis data yang besar. *Data mining* telah berkembang menjadi suatu bidang studi yang terintegrasi dalam area kecerdasan buatan dan rekayasa pengetahuan. Meskipun *data mining* berakar pada *machine learning* dan statistik, namun juga meluas ke bidang lain dalam ilmu komputer dan bidang lainnya seperti biologi, lingkungan, keuangan, jaringan, dan banyak lagi (Sudarsono et al., 2021).

Teknik-teknik data mining adalah sebagai berikut: (Rahayu et al., 2024)

a. Klasifikasi (*Classification*)

   Pada teknik ini digunakan untuk pengelompokan data ke dalam kelas atau kategori tertentu berdasarkan kesamaan atribut atau karakteristik. Contoh teknik klasifikasi ini mencakup penggunaan *Decision Tree*, *Artificial Neural Network*, dan berbagai algoritma klasifikasi lainnya.

b. Regresi (*Regression*)

   Untuk menganalisis dalam pengolahan data yang bertujuan untuk memperkirakan interaksi antara variabel bebas dan terikat, serta untuk memprediksi nilai kontinu dari variabel terikat berdasarkan pola interaksi baik yang linier maupun nonlinier.

c. Klastering (*Clustering*)

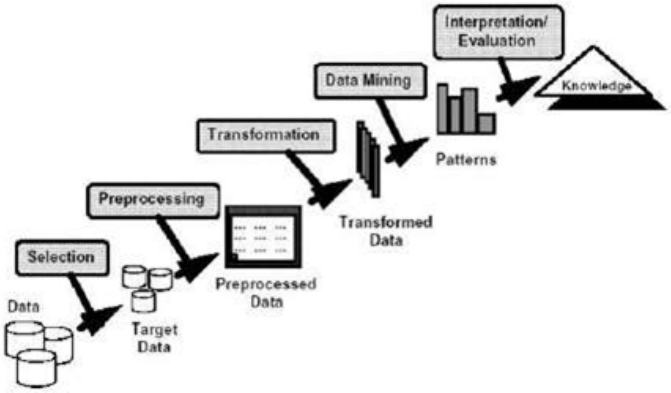
   Clustering digunakan untuk mengelompokkan data menjadi beberapa kelompok (*cluster*) berdasarkan kesamaan karakteristik yang dimiliki antar data. Contoh pada

teknik ini yang digunakan adalah *K-Means*, *K-Medoids*, *Hierarchical Clustering*, dan algoritma *clustering* lainnya.

d. Asosiasi (*Association*)

Teknik asosiasi digunakan untuk menemukan hubungan unik atau menarik yang tidak terlihat dalam kumpulan data yang besar. Hubungan yang tidak terlihat ini bisa dipresentasikan dalam bentuk aturan asosiasi atau kumpulan barang yang sering muncul. Contoh pada teknik ini yaitu *Apriori Algorithm*, *Eclat*, dan *FP-Growth*.

Tahapan dalam data mining adalah sebagai berikut: (Surahmat & Sutrisno, 2023)



Gambar 2. Tahapan Data Mining

Proses *cleaning* data memiliki tujuan menghapus data yang tidak konsisten serta menghapus atribut yang tidak sesuai dengan hipotesis penelitian.

1. Integrasi Data

   Tahapan ini digunakan untuk menggabungkan data dari beberapa sumber yang ada dan dijadikan satu sebagai data utama.

2. Transformasi Data

   Tahap *transformasi* data dilakukan dengan mengubah data ke dalam bentuk atau format yang sesuai dengan kebutuhan penelitian.

3. Aplikasi Teknik Data Mining

   Tahapan ini merupakan bagian dari proses *data mining* dan perlu diperhatikan bahwa dalam berbagai teknik yang telah ada dapat memastikan keberhasilan dalam pelaksanaan proses *data mining* tertentu.

4. Evaluasi Pola

   Pada tahap evaluasi pola yang telah ditemukan dapat digunakan untuk menemukan pola-pola lain yang memiliki keunikan tertentu serta prediksi lain yang memiliki nilai.

5. Presentasi Pola

Pada tahap presentasi ini, pola yang telah ditemukan dapat digunakan untuk menentukan tindakan terkait penelitian yang sedang dilakukan.

2.3 Algoritma C4.5

Algoritma C4.5 adalah metode yang digunakan untuk memprediksi data melalui teknik pohon keputusan. J. Ross Quinlan mengembangkan algoritma C4.5 sebagai penyempurnaan dari algoritma ID3 pada tahun 1996. Konsep utama algoritma ini adalah membangun struktur pohon keputusan dengan memilih peringkat tertinggi atau yang dapat disebut memiliki nilai gain tertinggi, dengan menggunakan nilai *entropy* atribut tersebut sebagai pusat klasifikasi atribut. Lalu, cabang-cabang pohon diperluas secara berulang hingga seluruh struktur pohon terbentuk (M. D. Wahyudi, 2023).

2.4 *Decision Tree*

Decision tree merupakan salah satu metode klasifikasi yang paling banyak digunakan karena hasilnya mudah dipahami oleh manusia. Metode ini membentuk model prediksi dengan menggunakan struktur pohon atau hierarki. Pemilihan atribut sebagai akar didasarkan pada nilai *Gain* tertinggi dari seluruh atribut yang tersedia (A. K. Wahyudi et al., 2022).

2.5 *Confusion Matrix*

Confusion Matrix memberikan gambaran umum tentang hasil prediksi dalam suatu masalah klasifikasi. *Confusion Matrix* digunakan untuk menganalisis seberapa baik model klasifikasi yang diterapkan dalam mengenali data dari berbagai kategori (Amillina & Qoiriah, 2021). Berikut merupakan dari *Confusion Matrix* :

Tabel 1. *Confusion Matrix*

		Prediksi Kelas	
		Yes	***No***
Kelas Sebenarnya	***Yes***	*TP*	*FN*
	No	*FP*	*TN*

Keterangan :

Kelas sebenarnya (*yes*) : kelas aktual yang memiliki nilai *yes*.

Kelas sebenarnya (*no*) : kelas aktual yang memiliki nilai *no*.
Prediksi kelas (*yes*) : kelas yang diprediksi dengan nilai *yes*.
Prediksi kelas (*no*) : kelas yang diprediksi dengan nilai *no*.
TP (*True Positive*) : prediksi yang positif dan kelas yang sebenarnya positif.
TN (*True Negative*) : prediksi negatif dengan kelas yang sebenarnya negatif.
FP (*False Positive*) : prediksi positif dengan kelas yang sebenarnya positif.
FN (*False Negative*) : prediksi negatif dengan kelas yang sebenarnya positif.

True Positive dan *True Negative* menandakan bahwa model klasifikasi mendefinisikan secara akurat, sedangkan *False Positive* atau *False Negative* menandakan bahwa model tersebut mendefinisikan secara salah. Penilaian *Accuracy, Precision,* dan *Recall* dapat dilakukan dengan menggunakan *Confusion matrix* sebagai berikut:

a. *Accuracy* merupakan mengacu pada tingkat keakuratan hasil prediksi. Untuk menentukan keakuratan, harus digunakan satu rumus.

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP}$$

b. *Precision* merupakan tingkat keakuratan informasi yang diinginkan pengguna dibandingkan dengan hasil prediksi yang dihasilkan oleh sistem tersebut. Dengan kata lain, ini adalah persentase dari prediksi yang tepat. Presisi dapat ditentukan melalui suatu rumus.

$$Precision = \frac{TP}{TP + FP}$$

c. *Recall* merupakan tingkat efektivitas sistem dalam menemukan hasil kasus positif yang telah diidentifikasi dengan tepat. Recall dapat dihitung menggunakan rumus.

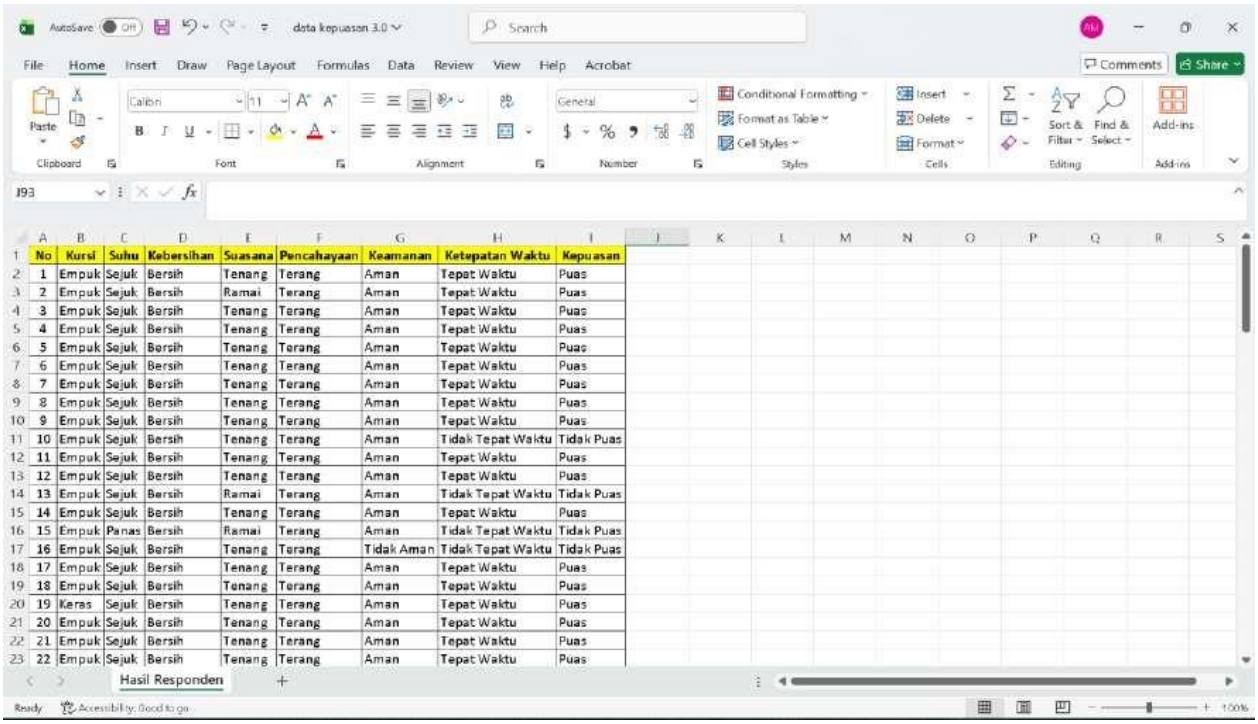
$$Recall = \frac{TP}{TP + FN}$$

2.6 RapidMiner

RapidMiner merupakan aplikasi yang berfungsi untuk mengelola serta menganalisis data. Aplikasi ini menyediakan solusi untuk melakukan analisis terhadap *data mining*, *text mining*, serta analisis prediktif (Lestari & Mulyawan, 2023). Perangkat lunak *RapidMiner* menyediakan antarmuka pengguna yang intuitif untuk analisis data, memungkinkan pengguna memproses, memodelkan, dan menampilkan hasil data dengan cara yang efisien tanpa memerlukan keterampilan pemrograman tingkat lanjut.

3. HASIL DAN PEMBAHASAN

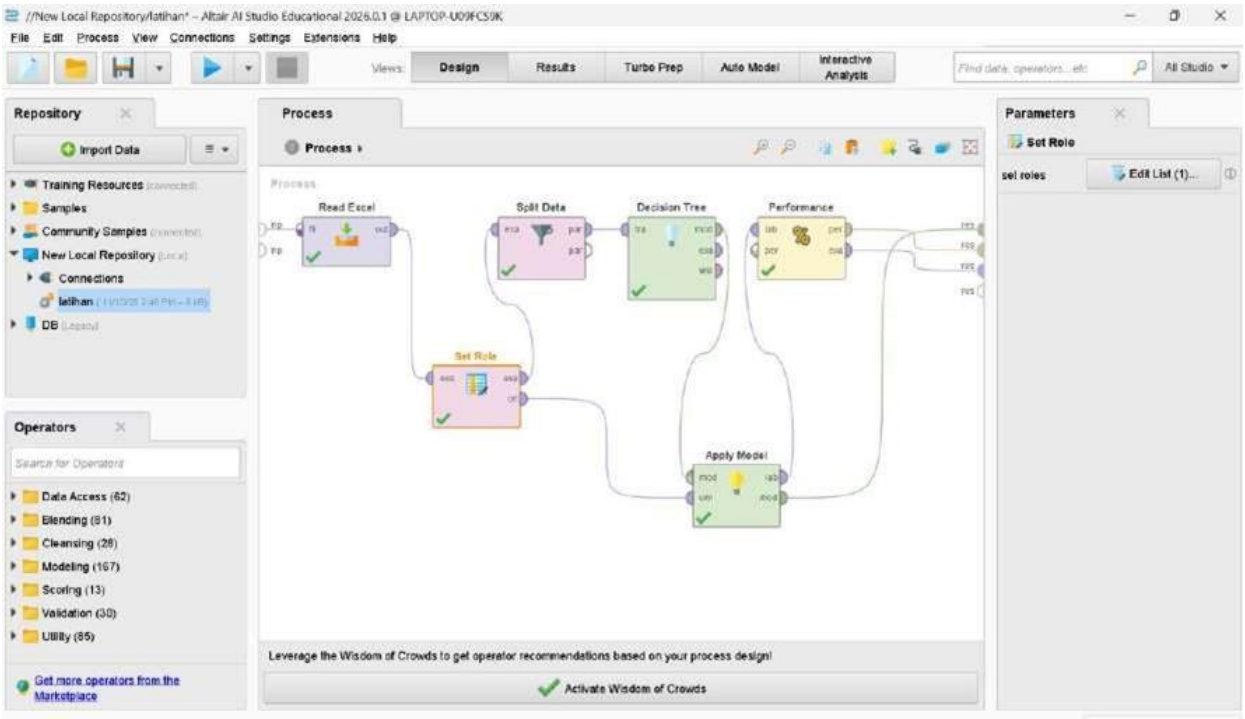
A. Data kuesioner penelitian tentang pengguna MRT dibuat dalam bentuk *file* Excel melalui *Google Forms*. Responden yang terlibat berjumlah 100 orang, khususnya kelompok generasi Z, dan kuesioner mencakup delapan atribut seperti yang terlihat pada gambar berikut:



No	Kursi	Suhu	Kebersihan	Suasana	Pencahayaan	Keamanan	Ketepatan Waktu	Kepuasan
1	Empuk	Sejuk	Bersih	Tenang	Terang	Aman	Tepat Waktu	Puas
2	Empuk	Sejuk	Bersih	Ramai	Terang	Aman	Tepat Waktu	Puas
3	Empuk	Sejuk	Bersih	Tenang	Terang	Aman	Tepat Waktu	Puas
4	Empuk	Sejuk	Bersih	Tenang	Terang	Aman	Tepat Waktu	Puas
5	Empuk	Sejuk	Bersih	Tenang	Terang	Aman	Tepat Waktu	Puas
6	Empuk	Sejuk	Bersih	Tenang	Terang	Aman	Tepat Waktu	Puas
7	Empuk	Sejuk	Bersih	Tenang	Terang	Aman	Tepat Waktu	Puas
8	Empuk	Sejuk	Bersih	Tenang	Terang	Aman	Tepat Waktu	Puas
9	Empuk	Sejuk	Bersih	Tenang	Terang	Aman	Tepat Waktu	Puas
10	Empuk	Sejuk	Bersih	Tenang	Terang	Aman	Tidak Tepat Waktu	Tidak Puas
11	Empuk	Sejuk	Bersih	Tenang	Terang	Aman	Tepat Waktu	Puas
12	Empuk	Sejuk	Bersih	Tenang	Terang	Aman	Tepat Waktu	Puas
13	Empuk	Sejuk	Bersih	Ramai	Terang	Aman	Tidak Tepat Waktu	Tidak Puas
14	Empuk	Sejuk	Bersih	Tenang	Terang	Aman	Tepat Waktu	Puas
15	Empuk	Panas	Bersih	Ramai	Terang	Aman	Tidak Tepat Waktu	Tidak Puas
16	Empuk	Sejuk	Bersih	Tenang	Terang	Tidak Aman	Tidak Tepat Waktu	Tidak Puas
17	Empuk	Sejuk	Bersih	Tenang	Terang	Aman	Tepat Waktu	Puas
18	Empuk	Sejuk	Bersih	Tenang	Terang	Aman	Tepat Waktu	Puas
19	Keras	Sejuk	Bersih	Tenang	Terang	Aman	Tepat Waktu	Puas
20	Empuk	Sejuk	Bersih	Tenang	Terang	Aman	Tepat Waktu	Puas
21	Empuk	Sejuk	Bersih	Tenang	Terang	Aman	Tepat Waktu	Puas
22	Empuk	Sejuk	Bersih	Tenang	Terang	Aman	Tepat Waktu	Puas

Gambar 3. Data Kuesioner

B. Pengolahan data: Dalam pengujian ini, *RapidMiner* digunakan untuk mengumpulkan hasil *accuracy* yang diuji menggunakan data kuesioner ini. Berikut adalah proses *preprocessing*:



Gambar 4. Proses Preprocessing

Setiap operator memiliki tugas yang berbeda dan saling mendukung agar bisa membuat model klasifikasi yang benar. Berikut penjelasan mengenai peran masing-masing operator:

1. *Read Excel*

   Operator ini berfungsi untuk mengimpor data kuesioner dari *file* Excel ke dalam lingkungan *RapidMiner*. Data yang dibaca kemudian dikonversi menjadi *ExampleSet* sebagai format standar yang digunakan dalam proses data mining. Operator ini memastikan bahwa seluruh atribut dan nilai pada dataset dapat diproses oleh operator selanjutnya.

2. *Set Role*

   Operator *Set Role* digunakan untuk menetapkan peran tertentu pada atribut, terutama dalam menentukan atribut mana yang menjadi *label* atau target klasifikasi. Pada penelitian ini, atribut "Kepuasan" ditetapkan sebagai *label*, sementara atribut lainnya diperlakukan sebagai *predictor*. Penetapan peran ini sangat penting agar algoritma dapat mengidentifikasi variabel target dengan benar.

3. *Split Data*

   Operator ini digunakan untuk membagi dataset menjadi data latih (*training set*) dan data uji (*testing set*). Pembagian data dilakukan agar model dapat dilatih menggunakan sebagian data, kemudian dievaluasi menggunakan data yang belum pernah dilihat oleh model. Pada penelitian ini, metode *holdout* digunakan untuk menghasilkan evaluasi model yang objektif.

4. *Decision Tree (C4.5)*

   Operator Decision Tree berfungsi untuk membangun model klasifikasi berbasis pohon keputusan menggunakan algoritma C4.5. Operator ini melakukan pemilihan atribut terbaik berdasarkan nilai *Gain Ratio* untuk membentuk struktur pohon. Model yang dihasilkan menggambarkan hubungan antara variabel-variabel dalam menentukan tingkat kepuasan pengguna MRT.

5. *Apply Model*

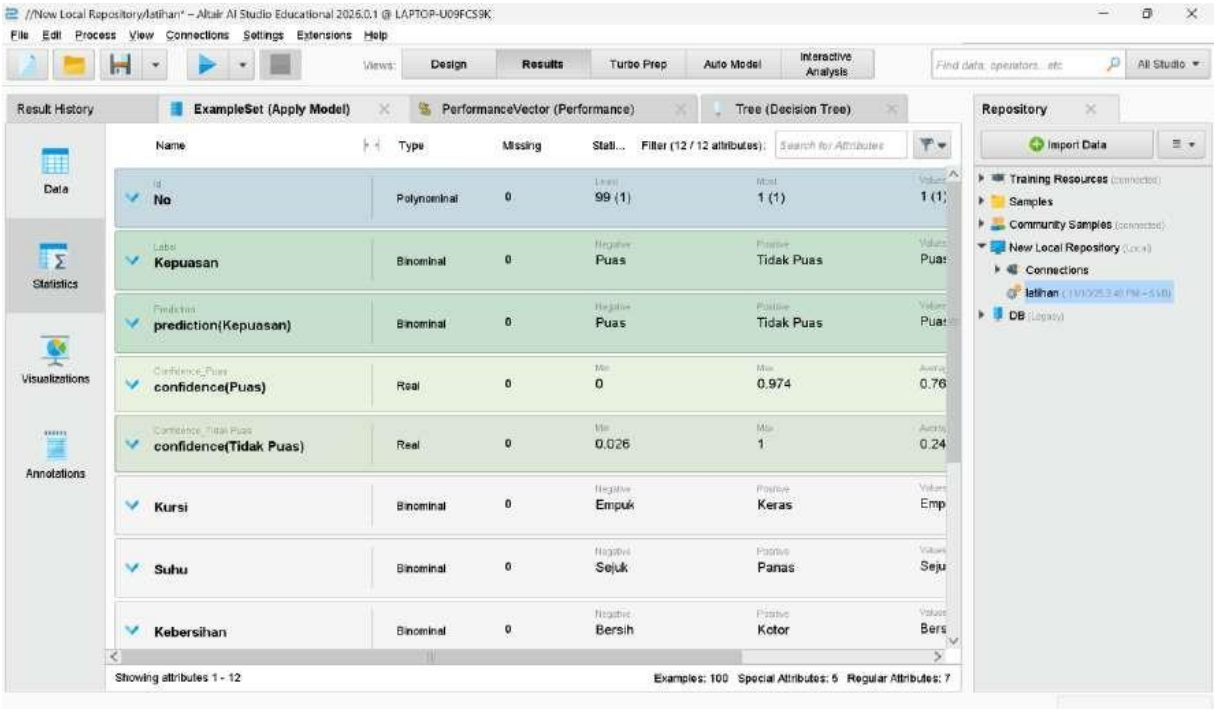
   Operator ini digunakan untuk menerapkan model *decision tree* yang telah dilatih pada data uji. Hasil prediksi dari proses ini menunjukkan kelas yang diprediksi (Puas atau Tidak Puas) serta skor kepercayaannya (*confidence value*). Langkah ini menjadi dasar untuk menilai performa model.

6. *Performance*

   Operator *Performance* digunakan untuk mengukur kinerja model berdasarkan hasil prediksi yang dihasilkan oleh operator *Apply Model*. Beberapa metrik evaluasi utama yang dihitung meliputi *Accuracy, Precision,* dan *Recall*, serta penyajian *confusion matrix*. Operator ini memberikan gambaran yang jelas mengenai tingkat efektivitas model C4.5 dalam mengklasifikasikan kepuasan pengguna MRT.

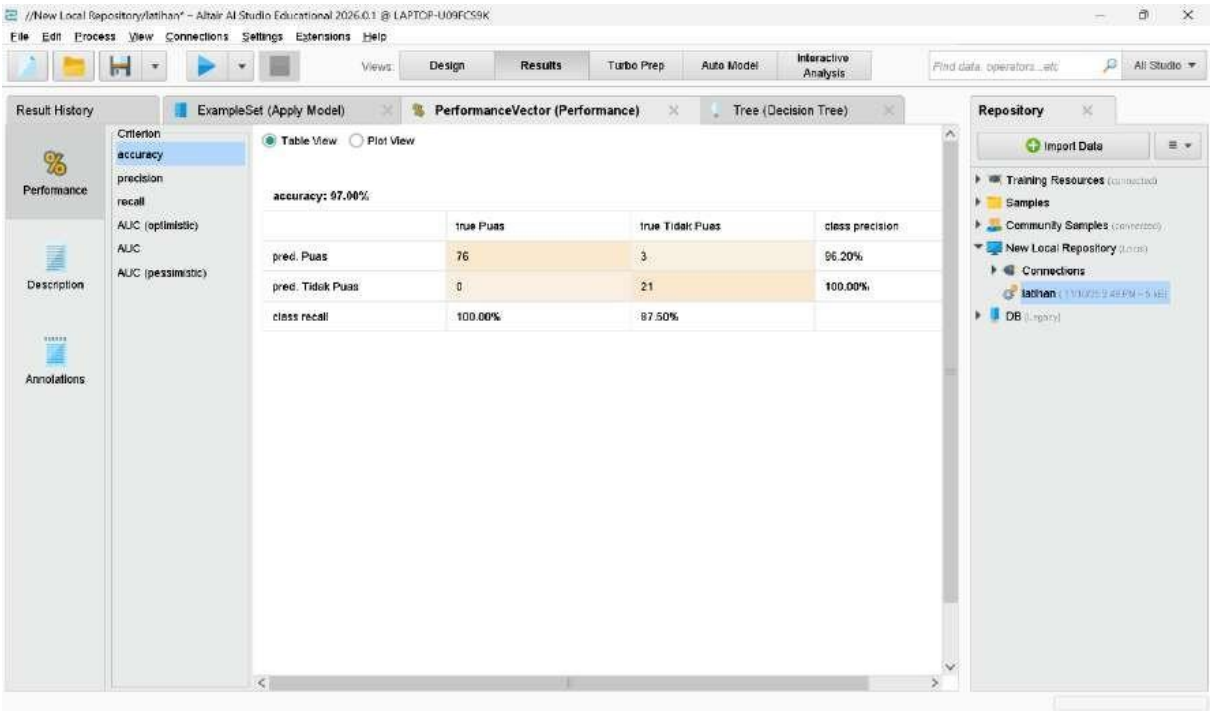
C. Eksekusi Proses: Setelah semua operator terhubung dengan benar, klik tombol *Run* untuk memulai proses. *RapidMiner* akan menampilkan struktur pohon keputusan.

a. Hasil dari data *preprocessing*



Gambar 5. Hasil Data Preprocessing

b. Berdasarkan hasil analisis data menggunakan aplikasi *RapidMiner*, tingkat *accuracy* yang berhasil dicapai oleh algoritma C4.5 adalah 97.00%.



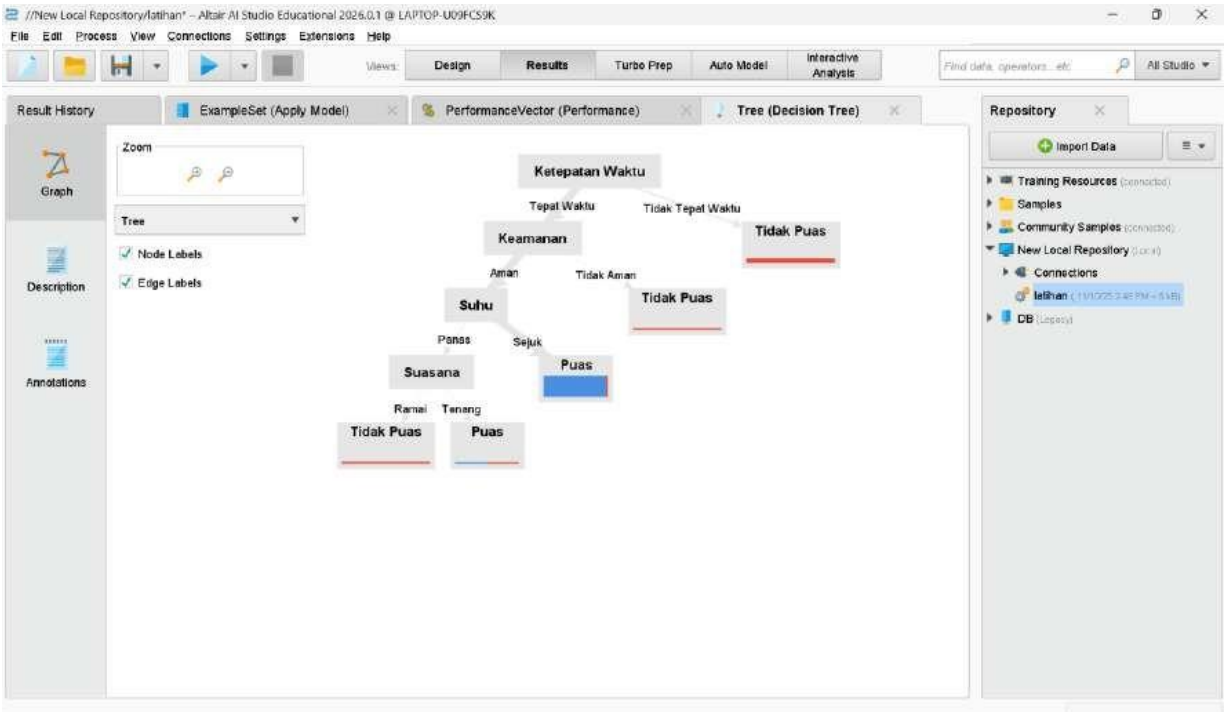
Gambar 6. Tingkat Akurasi

Hasil dari proses data menggunakan algoritma C4.5 dapat dilihat pada tabel berikut:

Tabel 2. Evaluasi hasil pengolahan data

METODE	***Accuracy***	***Precision***	***Recal***
Algoritma C4.5	97.00%	100.00%	87.50%

c. Hasil dari pengolahan data tersebut telah menghasilkan sebuah model pohon keputusan.

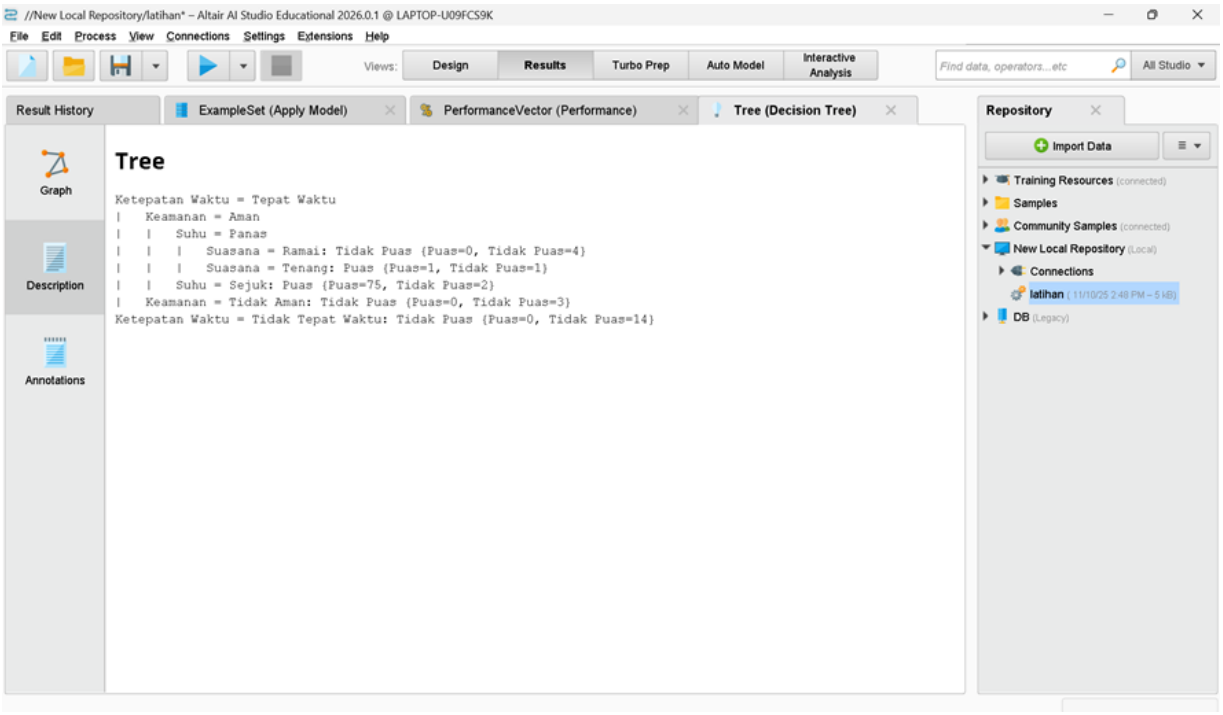


Gambar 7. Pohon Keputusan

Berdasarkan hasil analisis klasifikasi menggunakan algoritma C4.5, diperoleh struktur *decision tree* yang menunjukkan bahwa atribut Ketepatan Waktu berperan sebagai akar keputusan utama dalam menentukan tingkat kepuasan pengguna MRT dari kategori Generasi Z. Hasil ini mengindikasikan bahwa ketepatan waktu menjadi faktor paling berpengaruh terhadap persepsi kepuasan pengguna. Apabila layanan MRT dinilai tidak tepat waktu, mayoritas pengguna menunjukkan respon tidak puas terhadap layanan yang diberikan. Sebaliknya, apabila layanan dianggap tepat waktu, maka faktor berikutnya yang berperan adalah Keamanan.

Pengguna yang merasa tidak aman cenderung memberikan penilaian tidak puas, sedangkan pengguna yang merasa aman selanjutnya dipengaruhi oleh faktor Suhu di dalam lingkungan MRT. Pada kondisi suhu sejuk, pengguna menunjukkan kecenderungan puas, sedangkan pada kondisi suhu panas, kepuasan pengguna dipengaruhi oleh Suasana di dalam kereta. Apabila suasana tergolong tenang, pengguna cenderung merasa puas, namun jika suasana ramai, maka tingkat kepuasan menurun.

Secara keseluruhan, hasil pemodelan ini menegaskan bahwa ketepatan waktu merupakan determinan utama dalam pembentukan kepuasan pengguna MRT di kalangan Generasi Z, diikuti oleh faktor keamanan, kenyamanan suhu, serta suasana lingkungan. Dengan demikian, peningkatan kualitas layanan pada aspek-aspek tersebut diperkirakan dapat memberikan kontribusi signifikan terhadap peningkatan tingkat kepuasan pengguna MRT.



Gambar 8. Rule Yang Dihasilkan Dari Pohon Keputusan

KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa menggunakan metode data mining dengan algoritma C4.5 berhasil mengklasifikasikan tingkat kepuasan pengguna MRT dari kelompok Generasi Z secara efektif. Struktur decision tree yang dihasilkan menunjukkan bahwa ketepatan waktu adalah faktor utama dalam menentukan kepuasan pengguna MRT, diikuti oleh faktor keamanan, suhu, dan suasana di dalam kereta. Artinya, pengguna Generasi Z paling mengutamakan konsistensi waktu keberangkatan dan kedatangan MRT sebagai penentu utama kepuasan.

Untuk penelitian berikutnya, disarankan untuk memperluas cakupan kelompok usia agar mendapatkan hasil yang lebih lebih representatif. Selain itu, penambahan variabel seperti kebersihan, aksesibilitas, dan efisiensi tarif diharapkan dapat meningkatkan analisis yang lebih dalam terhadap faktor-faktor yang memengaruhi kepuasan pengguna. Penelitian selanjutnya juga bisa membandingkan performa antara algoritma C4.5 dengan algoritma klasifikasi lainnya untuk menilai tingkat akurasi serta efektivitas metode dalam menganalisis kepuasan pengguna transportasi umum.

DAFTAR PUSTAKA

Amillina, I., & Qoiriah, A. (2021). Penerapan Algoritma Naïve Bayes dalam Klasifikasi Tingkat Kepuasan Siswa terhadap Pembelajaran Daring. *JIFTI - Jurnal Ilmiah Teknologi Informasi Dan Robotika*, *3*, 16–23. https://doi.org/https://doi.org/10.33005/jifti.v3i2.59

Lestari, P. D., & Mulyawan. (2023). DATAMINING PADA PENJUALAN AIR BERSIH DI SPAM AKIDAH MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING MENGGUNAKAN RAPIDMINER. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *7*(1), 412–416. https://doi.org/https://dx.doi.org/10.36040/jati.v7i1.6315

Pribadi, M. N. N., & Ernawati, L. (2024). Perbandingan Hasil Penerapan Algoritma Klasifikasi Dan Natural Language Processing Terhadap Data Kepuasan Pengguna Layanan Transportasi Umum MRT Jakarta. *Informatik : Jurnal Ilmu Komputer*, *20*(3), 102–111. https://doi.org/10.52958/iftk.v20i3.10502

Purnandika, R. A., & Septiana, H. (2023). Kualitas Layanan Terhadap Kepuasan Pelanggan pada Pengguna Transportasi Umum di Jakarta. *Jurnal Ilmiah Manajemen Ekonomi Dan Akuntansi (JIMEA)*, *1*(2), 25–32. https://doi.org/10.62017/jimea.v1i2.226

Putri, N. I. P. (2022). Pengaruh Transportasi Umum Terhadap Perkembangan Antar Wilayah Administrasi di Jakarta. *Jurnal Manajemen Bisnis Transportasi Dan Logistik*, *8*(1), 63. https://doi.org/10.54324/j.mbtl.v8i1.1249

Rahayu, P. W., Sudipa, I. G. I., Suryani, Surachman, A., Ridwan, A., Darmawiguna, I. G. M., Sutoyo, M. N., Harlina, S., & Maysanjaya, I. M. D. (2024). *Buku ajar data mining* (Efitra (ed.); Issue January). PT. Sonpedia Publishing Indonesia. https://books.google.co.id/books?hl=id&lr=&id=vCruEAAAQBAJ&oi=fnd&pg=PA83&dq=BUKU+AJAR+DATA+MINING&ots=TpLL6QpOqB&sig=L-U7CXxD6yVCNGdgAHQio6w89-E&redir_esc=y#v=onepage&q=BUKU AJAR DATA MINING&f=false

Statistik, B. P. (2025). Perkembangan Transportasi DKI Jakarta Desember 2024. *Badan Pusat Statistic Jakarta Utara*, *13*. https://jakutkota.bps.go.id/id/pressrelease/2025/02/03/452/perkembangan-transportasi-dki-jakarta-desember-2023.html

Sudarsono, B. G., Leo, M. I., Santoso, A., & Hendrawan, F. (2021). Analisis Data Mining Data Netflix Menggunakan Aplikasi Rapid Miner Analysis Data Mining Netflix Data Using The Rapid Miner Application. *Journal of Business and Audit Information Systems*, *4*(1), 13–21. https://doi.org/http://dx.doi.org/10.30813/jbase.v4i1.2729 Hasil

Surahmat, A., & Sutrisno, M. (2023). Analisis Kepuasan Pelanggan Dalam Industri Teknologi Menggunakan Algoritma C4.5. *Just IT : Jurnal Sistem Informasi, Teknologi Informasi Dan Komputer*, *13*(2), 75–79. https://doi.org/https://doi.org/10.24853/justit.13.2.75-79

Wahyudi, A. K., Azizah, N., & Saputro, H. (2022). DATA MINING KLASIFIKASI KEPRIBADIAN SISWA SMP NEGERI 5 JEPARA MENGGUNAKAN METODE DECISION TREE ALGORITMA C4 . 5. *Journal of Information System and Computer*, *2*(2), 8–13. https://doi.org/https://doi.org/10.34001/jister.v2i2.392

Wahyudi, M. D. (2023). PENERAPAN DATA MINING DENGAN ALGORITMA C4.5 DALAM PREDIKSI PENJUALAN BUKU. *Jurnal Teknorama (Informatika Dan Teknologi El Rahma)*, *1*(1), 1–6. https://doi.org/https://doi.org/10.59812/teknorama.v1i1.1