

BAB III

METODOLOGI PENELITIAN

3.1. Jenis Penelitian

Jenis penelitian yang digunakan merupakan kombinasi antara pendekatan kualitatif dan kuantitatif. Pendekatan kuantitatif digunakan untuk menggambarkan persepsi pengunjung terhadap layanan Cafe Teman Nongkrong, seperti rasa, harga menu, pelayanan, kebersihan, dan kecepatan penyajian yang diperoleh melalui kuesioner berbasis Google Form. Sementara itu, pendekatan kualitatif digunakan untuk melihat seberapa besar kualitas metode yang diterapkan, yaitu dengan mengukur hasil evaluasi model Decision Tree melalui nilai akurasi, presisi, recall, dan F1-Score, sehingga kinerja metode dapat dinilai secara objektif.

Tahap selanjutnya adalah pengolahan data, yang meliputi proses pembersihan dan seleksi data untuk menghilangkan data yang tidak lengkap, duplikat, atau tidak valid. Data yang telah bersih kemudian disiapkan dan diubah ke dalam format yang sesuai untuk proses analisis. Setelah itu, data dibagi menjadi data pelatihan dan data pengujian agar model yang dibangun dapat dilatih dan dievaluasi secara objektif.

Pada tahap pemodelan, data pelatihan diproses menggunakan algoritma Decision Tree dengan bantuan perangkat lunak RapidMiner. Algoritma ini bekerja dengan menghitung nilai entropy dan information gain dari setiap atribut untuk menentukan struktur pohon keputusan. Model yang dihasilkan kemudian digunakan untuk melakukan klasifikasi tingkat kepuasan pengunjung.

Tahap akhir dari arsitektur sistem adalah hasil klasifikasi dan analisis, di mana model menghasilkan output berupa kategori kepuasan pengunjung, yaitu puas atau tidak puas. Hasil tersebut dianalisis untuk mengetahui faktor-faktor yang paling berpengaruh terhadap kepuasan berdasarkan urutan nilai gain. Informasi yang diperoleh dari sistem ini dapat dijadikan dasar dalam pengambilan keputusan untuk meningkatkan kualitas layanan di Cafe Teman Nongkrong secara lebih terarah dan berbasis data.

3.2. Langkah-Langkah Pengolahan Data

Pengolahan data dilakukan melalui serangkaian langkah yang terstruktur agar proses analisis dapat berjalan dengan baik dan menghasilkan informasi yang akurat. Setiap tahapan memiliki peran penting, mulai dari pengumpulan data, pembersihan data, hingga pemodelan dan evaluasi, sehingga data yang digunakan benar-benar layak untuk dianalisis. Adanya langkah-langkah pengolahan data membantu meminimalkan kesalahan, menghindari data yang tidak valid, serta memastikan bahwa proses klasifikasi dilakukan secara sistematis. Dengan mengikuti tahapan pengolahan data secara runtut, hasil analisis yang diperoleh menjadi lebih objektif, dapat dipercaya, dan mampu mendukung pencapaian tujuan penelitian secara optimal.

3.2.1. Analisis Data

Analisis data merupakan tahapan awal yang dilakukan untuk menentukan variabel-variabel yang relevan dan sesuai dengan tujuan penelitian. Pada tahap ini, data yang tersedia ditelaah secara menyeluruh untuk memastikan setiap variabel memiliki peran dalam mendukung proses klasifikasi dan analisis tingkat kepuasan.

Penentuan variabel dilakukan secara selektif agar data yang digunakan bersifat terarah, tidak berlebihan, dan mampu menggambarkan kondisi yang diteliti secara akurat. Variabel-variabel yang digunakan selanjutnya akan dipaparkan secara rinci dalam bentuk tabel.

Tabel 3. 1. Variabel Penelitian

No	Variabel	Pernyataan
1	Rasa	Saya merasa rasa makanan/minuman yang disajikan sangat sesuai dengan selera saya.
		Makanan/minuman memiliki cita rasa yang konsisten setiap kali saya berkunjung.
		Bahan yang digunakan terasa segar dan berkualitas.
		Porsi makanan/minuman sesuai dengan rasa yang dihasilkan.
		Secara keseluruhan, saya puas dengan rasa makanan/minuman yang disajikan.
2	Harga Menu	Harga menu sesuai dengan kualitas makanan/minuman yang saya terima.
		Harga yang ditawarkan terasa wajar dengan suasana dan fasilitas kafe.
		Menu memiliki pilihan harga yang cukup variatif untuk berbagai kalangan.
		Saya merasa harga menu tidak terlalu mahal untuk ukuran kafe seperti ini.
		Secara keseluruhan, saya menilai harga menu berada pada kategori yang sesuai.
3	Pelayanan	Staf kafe melayani dengan sikap ramah dan sopan.
		Pelayanan terasa cepat dan tidak membuat saya menunggu lama.

		Staf kafe tanggap ketika saya membutuhkan bantuan atau informasi.
		Interaksi staf memberikan pengalaman yang menyenangkan bagi saya.
		Secara keseluruhan, saya puas dengan pelayanan yang diberikan oleh staf.
4	Kebersihan	Area meja dan tempat duduk selalu dalam kondisi bersih ketika saya datang.
		Lantai dan area sekitar kafe tampak terawat dan bebas dari kotoran.
		Peralatan makan/minum (gelas, piring, sendok) terlihat bersih.
		Area pelayanan seperti kasir dan tempat pengambilan pesanan terjaga kebersihannya.
		Secara keseluruhan, kafe memiliki tingkat kebersihan yang memuaskan.
5	Kecepatan Penyajian	Pesanan saya disajikan dalam waktu yang tidak terlalu lama.
		Proses pemesanan berlangsung cepat tanpa hambatan.
		Staf terlihat cekatan dalam menyiapkan pesanan pelanggan.
		Waktu penyajian sesuai dengan ekspektasi saya terhadap kafe ini.
		Secara keseluruhan, saya menilai kecepatan penyajian di kafe ini baik.

Tabel di atas merupakan daftar pernyataan kuesioner yang digunakan sebagai instrumen pengumpulan data untuk mengukur tingkat kepuasan pengunjung. Pernyataan-pernyataan tersebut disusun berdasarkan beberapa variabel utama yang merepresentasikan pengalaman pengunjung, mulai dari aspek rasa, harga menu,

pelayanan, kebersihan, hingga kecepatan penyajian. Setiap variabel diwakili oleh lima pernyataan yang dirancang untuk menangkap persepsi responden secara lebih mendalam dan terukur. Melalui pernyataan-pernyataan tersebut, data yang diperoleh diharapkan mampu menggambarkan kondisi kepuasan pengunjung secara objektif dan menjadi dasar dalam proses analisis serta klasifikasi tingkat kepuasan.

3.2.2. Pengumpulan Data

Pengumpulan data dilakukan dengan menggunakan kuesioner berbasis Google Form yang disebarakan kepada pengunjung Cafe Teman Nongkrong. Media ini dipilih karena memudahkan responden dalam memberikan penilaian serta memungkinkan data terkumpul secara otomatis dan terstruktur. Dari proses pengumpulan tersebut, diperoleh sebanyak 50 data responden yang selanjutnya digunakan sebagai bahan analisis. Data yang terkumpul menjadi dasar dalam proses pengolahan dan pemodelan untuk mengklasifikasikan tingkat kepuasan pengunjung secara sistematis.

Tabel 3. 2. Data Hasil Kuesioner

Nama Lengkap	Rasa	Harga Menu	Pelayanan	Kebersihan	Kecepatan Penyajian	Kategori
M* E*	Kurang Enak	Murah	Ramah	Bersih	Cepat	Puas
T*** R*** Y***	Enak	Mahal	Ramah	Bersih	Cepat	Puas
N**** A*****	Enak	Murah	Kurang Ramah	Bersih	Cepat	Puas
N***** A*****	Enak	Murah	Ramah	Kurang Bersih	Cepat	Puas
D** P*****	Enak	Murah	Ramah	Bersih	Lambat	Puas
I**** K*****	Enak	Murah	Ramah	Kurang Bersih	Cepat	Puas
M***** Y****	Enak	Mahal	Ramah	Bersih	Lambat	Puas

N***** S**	Kurang Enak	Murah	Ramah	Bersih	Cepat	Puas
P**** A***** R*****	Enak	Mahal	Ramah	Bersih	Cepat	Puas
N*****	Enak	Mahal	Ramah	Kurang Bersih	Cepat	Puas
Z**** D** A*****	Enak	Mahal	Ramah	Bersih	Cepat	Puas
S***** A**** T*****	Enak	Mahal	Kurang Ramah	Kurang Bersih	Lambat	Tidak Puas
S** Z***** J*****	Kurang Enak	Murah	Kurang Ramah	Kurang Bersih	Lambat	Tidak Puas
A***	Kurang Enak	Mahal	Ramah	Kurang Bersih	Lambat	Tidak Puas
R***** H*****	Kurang Enak	Mahal	Kurang Ramah	Bersih	Lambat	Tidak Puas
B***** S***** S*****	Kurang Enak	Mahal	Kurang Ramah	Kurang Bersih	Cepat	Tidak Puas
I**** R*****	Kurang Enak	Mahal	Kurang Ramah	Bersih	Lambat	Tidak Puas
S***** R****	Kurang Enak	Mahal	Ramah	Kurang Bersih	Lambat	Tidak Puas
R****	Kurang Enak	Murah	Kurang Ramah	Kurang Bersih	Lambat	Tidak Puas
A** L*****	Enak	Mahal	Kurang Ramah	Kurang Bersih	Lambat	Tidak Puas
S** S*****	Kurang Enak	Murah	Kurang Ramah	Kurang Bersih	Lambat	Tidak Puas
S** R*****	Kurang Enak	Mahal	Ramah	Kurang Bersih	Lambat	Tidak Puas
B** S*****	Kurang Enak	Mahal	Kurang Ramah	Bersih	Lambat	Tidak Puas
A** K*****	Kurang Enak	Mahal	Kurang Ramah	Kurang Bersih	Cepat	Tidak Puas
A** H*****	Kurang Enak	Mahal	Kurang Ramah	Bersih	Lambat	Tidak Puas
T**** A*****	Kurang Enak	Mahal	Kurang Ramah	Bersih	Lambat	Tidak Puas
A***** N*****	Kurang Enak	Mahal	Kurang Ramah	Bersih	Lambat	Tidak Puas
D** N*****	Kurang Enak	Mahal	Ramah	Kurang Bersih	Lambat	Tidak Puas
D**** A*****	Kurang Enak	Murah	Kurang Ramah	Kurang Bersih	Lambat	Tidak Puas
M***** S*****	Enak	Mahal	Kurang Ramah	Kurang Bersih	Lambat	Tidak Puas
A**** S**** D***	Kurang Enak	Murah	Kurang Ramah	Kurang Bersih	Lambat	Tidak Puas
S***** S*****	Kurang Enak	Mahal	Ramah	Kurang Bersih	Lambat	Tidak Puas
K***** S***	Enak	Murah	Ramah	Bersih	Cepat	Puas
R**** W*****	Enak	Murah	Ramah	Bersih	Cepat	Puas
H**** S*****	Enak	Murah	Ramah	Bersih	Cepat	Puas

R**** R*****	Enak	Murah	Ramah	Bersih	Cepat	Puas
R**** A*****	Enak	Murah	Ramah	Bersih	Cepat	Puas
C***** T***** L***** R***	Enak	Murah	Ramah	Bersih	Cepat	Puas
N***** H****	Enak	Murah	Ramah	Bersih	Cepat	Puas
D*** T*****	Enak	Murah	Ramah	Bersih	Cepat	Puas
S** W*****	Enak	Murah	Ramah	Bersih	Cepat	Puas
H**** Z***** R*****	Enak	Murah	Ramah	Bersih	Cepat	Puas
P**** H*****	Enak	Murah	Ramah	Bersih	Cepat	Puas
A***** R*****	Enak	Murah	Ramah	Bersih	Cepat	Puas
A**** P*****	Enak	Murah	Ramah	Bersih	Cepat	Puas
A***** P*****	Enak	Murah	Ramah	Bersih	Cepat	Puas
L** H*****	Enak	Murah	Ramah	Bersih	Cepat	Puas
W**** A*****	Enak	Murah	Ramah	Bersih	Cepat	Puas
D*** N*****	Kurang Enak	Mahal	Kurang Ramah	Kurang Bersih	Lambat	Tidak Puas
D*****	Kurang Enak	Mahal	Kurang Ramah	Kurang Bersih	Lambat	Tidak Puas

Tabel di atas merupakan data yang diperoleh dari hasil pengisian kuesioner oleh responden dengan jumlah keseluruhan sebanyak 50 data. Data tersebut berisi informasi penilaian pengunjung terhadap beberapa aspek, yaitu rasa, harga menu, pelayanan, kebersihan, dan kecepatan penyajian, serta kategori tingkat kepuasan yang dihasilkan. Seluruh data telah diolah terlebih dahulu menggunakan aplikasi Microsoft Excel sehingga tersusun secara rapi dan terstruktur seperti yang ditampilkan pada tabel. Namun demikian, data tersebut masih belum dapat langsung digunakan untuk proses analisis lebih lanjut karena berpotensi mengandung ketidakkonsistenan, kesalahan penulisan, atau nilai yang tidak sesuai.

3.2.3. Pembersihan Data

Pembersihan data merupakan tahapan yang dilakukan untuk memastikan data yang digunakan berada dalam kondisi baik dan siap dianalisis. Proses ini mencakup pengecekan terhadap kesalahan penulisan, data yang tidak lengkap, duplikasi, serta

ketidaksesuaian nilai pada setiap atribut. Dengan melakukan pembersihan data, informasi yang digunakan menjadi lebih akurat dan konsisten sehingga dapat menghasilkan hasil analisis yang lebih valid dan dapat dipercaya.

Tabel 3. 3. Data Sampel Penelitian

Nama Lengkap	Rasa	Harga Menu	Pelayanan	Kebersihan	Kecepatan Penyajian	Kategori
A** H*****	Kurang Enak	Mahal	Kurang Ramah	Bersih	Lambat	Tidak Puas
A**** S**** D***	Kurang Enak	Murah	Kurang Ramah	Kurang Bersih	Lambat	Tidak Puas
A***** P*****	Enak	Murah	Ramah	Bersih	Cepat	Puas
A***** R*****	Enak	Murah	Ramah	Bersih	Cepat	Puas
A***	Kurang Enak	Mahal	Ramah	Kurang Bersih	Lambat	Tidak Puas
A*** K*****	Kurang Enak	Mahal	Kurang Ramah	Kurang Bersih	Cepat	Tidak Puas
A**** P*****	Enak	Murah	Ramah	Bersih	Cepat	Puas
A***** N*****	Kurang Enak	Mahal	Kurang Ramah	Bersih	Lambat	Tidak Puas
A** L*****	Enak	Mahal	Kurang Ramah	Kurang Bersih	Lambat	Tidak Puas
B***** S***** S*****	Kurang Enak	Mahal	Kurang Ramah	Kurang Bersih	Cepat	Tidak Puas
B*** S*****	Kurang Enak	Mahal	Kurang Ramah	Bersih	Lambat	Tidak Puas
C***** T***** L***** R***	Enak	Murah	Ramah	Bersih	Cepat	Puas
D*****	Kurang Enak	Mahal	Kurang Ramah	Kurang Bersih	Lambat	Tidak Puas
D*** T*****	Enak	Murah	Ramah	Bersih	Cepat	Puas
D*** N*****	Kurang Enak	Mahal	Kurang Ramah	Kurang Bersih	Lambat	Tidak Puas
D**** A*****	Kurang Enak	Murah	Kurang Ramah	Kurang Bersih	Lambat	Tidak Puas
D** N*****	Kurang Enak	Mahal	Ramah	Kurang Bersih	Lambat	Tidak Puas
D** P*****	Enak	Murah	Ramah	Bersih	Lambat	Puas
H**** Z***** R*****	Enak	Murah	Ramah	Bersih	Cepat	Puas
H**** S*****	Enak	Murah	Ramah	Bersih	Cepat	Puas
I**** K*****	Enak	Murah	Ramah	Kurang Bersih	Cepat	Puas
I**** R*****	Kurang Enak	Mahal	Kurang Ramah	Bersih	Lambat	Tidak Puas

K***** S***	Enak	Murah	Ramah	Bersih	Cepat	Puas
L** H*****	Enak	Murah	Ramah	Bersih	Cepat	Puas
M* E*	Kurang Enak	Murah	Ramah	Bersih	Cepat	Puas
M***** S*****	Enak	Mahal	Kurang Ramah	Kurang Bersih	Lambat	Tidak Puas
M***** Y****	Enak	Mahal	Ramah	Bersih	Lambat	Puas
N**** A*****	Enak	Murah	Kurang Ramah	Bersih	Cepat	Puas
N***** A*****	Enak	Murah	Ramah	Kurang Bersih	Cepat	Puas
N***** H****	Enak	Murah	Ramah	Bersih	Cepat	Puas
N*****	Enak	Mahal	Ramah	Kurang Bersih	Cepat	Puas
N***** S**	Kurang Enak	Murah	Ramah	Bersih	Cepat	Puas
P**** A***** R*****	Enak	Mahal	Ramah	Bersih	Cepat	Puas
P**** H*****	Enak	Murah	Ramah	Bersih	Cepat	Puas
R**** R*****	Enak	Murah	Ramah	Bersih	Cepat	Puas
R**** A*****	Enak	Murah	Ramah	Bersih	Cepat	Puas
R**** W*****	Enak	Murah	Ramah	Bersih	Cepat	Puas
R****	Kurang Enak	Murah	Kurang Ramah	Kurang Bersih	Lambat	Tidak Puas
R***** H*****	Kurang Enak	Mahal	Kurang Ramah	Bersih	Lambat	Tidak Puas
S***** R****	Kurang Enak	Mahal	Ramah	Kurang Bersih	Lambat	Tidak Puas
S***** A**** T*****	Enak	Mahal	Kurang Ramah	Kurang Bersih	Lambat	Tidak Puas
S*** Z***** J*****	Kurang Enak	Murah	Kurang Ramah	Kurang Bersih	Lambat	Tidak Puas
S***** S*****	Kurang Enak	Mahal	Ramah	Kurang Bersih	Lambat	Tidak Puas
S** R*****	Kurang Enak	Mahal	Ramah	Kurang Bersih	Lambat	Tidak Puas
S** W*****	Enak	Murah	Ramah	Bersih	Cepat	Puas
S*** S*****	Kurang Enak	Murah	Kurang Ramah	Kurang Bersih	Lambat	Tidak Puas
T**** A*****	Kurang Enak	Mahal	Kurang Ramah	Bersih	Lambat	Tidak Puas
T*** R*** Y***	Enak	Mahal	Ramah	Bersih	Cepat	Puas
W**** A*****	Enak	Murah	Ramah	Bersih	Cepat	Puas
Z**** D** A*****	Enak	Mahal	Ramah	Bersih	Cepat	Puas

Tabel di atas merupakan data yang telah melalui tahap pembersihan sehingga siap digunakan untuk proses analisis lebih lanjut. Pada tahap ini, data telah disusun kembali secara rapi dengan memperbaiki urutan data, penyeragaman penulisan,

serta memastikan setiap atribut terisi dengan nilai yang sesuai. Selain itu, kesalahan penulisan, inkonsistensi format, dan data yang berpotensi mengganggu proses pengolahan telah diperbaiki atau disesuaikan. Dengan kondisi data yang lebih terstruktur dan konsisten dibandingkan data sebelumnya, kualitas data menjadi lebih baik sehingga dapat mendukung proses pemodelan dan menghasilkan hasil analisis yang lebih akurat.

3.2.4. Pemilihan Model

Pemilihan model dilakukan dengan menentukan metode yang mampu mengklasifikasikan tingkat kepuasan secara jelas dan mudah dipahami, sehingga digunakan metode *Decision Tree*. Metode ini bekerja dengan membentuk struktur pohon keputusan yang menunjukkan hubungan antara atribut dan hasil klasifikasi berdasarkan proses perhitungan tertentu. *Decision Tree* memiliki keunggulan dalam menyajikan aturan keputusan yang sederhana serta mudah diinterpretasikan, sehingga dapat membantu mengetahui faktor-faktor yang paling berpengaruh terhadap tingkat kepuasan. Pada metode ini, proses pengolahan data dilakukan dengan memaparkan dan menerapkan rumus entropy, gain, dan \log_2 sebagai dasar perhitungan untuk menentukan atribut terbaik dalam pembentukan pohon keputusan.

1. Rumus Entropy (2 Kelas)

$$Entropy(S) = (-P_{Puas} \times \log_2(P_{Puas})) + (-P_{Tidak Puas} \times \log_2(P_{Tidak Puas}))$$

Keterangan:

- 1) Entropy (S) : nilai entropy pada dataset S
- 2) P_{Puas} : probabilitas data dengan kelas Puas

- 3) $P_{Tidak\ Puas}$: probabilitas data dengan kelas Tidak Puas
- 4) \log_2 : logaritma basis 2
- 5) Tanda (-) : digunakan agar hasil entropy bernilai positif
- 6) Info : Urutan perhitungan probabilitas kelas pada rumus entropy, baik dimulai dari $P_{Tidak\ Puas}$ kemudian P_{Puas} maupun sebaliknya, tidak memengaruhi nilai entropy yang dihasilkan karena entropy merupakan hasil penjumlahan dari masing-masing kelas.

2. Rumus Logaritma Baris 2 (Log2)

$$\log_2(x) = \frac{\ln(x)}{\ln(2)}$$

Keterangan:

- 1) $\log_2(x)$: logaritma bilangan x dengan basis 2
- 2) x : nilai probabilitas ($0 < x \leq 1$)
- 3) $\ln(x)$: logaritma natural
- 4) $\ln(2)$: logaritma natural dari angka 2
- 5) Info : Rumus ini digunakan jika kalkulator tidak menyediakan \log_2 secara langsung

3. Rumus Information Gain

$$Gain(S, A) = Entropy(S) - \left[\left(\frac{|S_1|}{|S|} \times Entropy(S_1) \right) + \left(\frac{|S_2|}{|S|} \times Entropy(S_2) \right) \right]$$

Keterangan:

- 1) Gain(S, A) : nilai information gain dari atribut A

- 2) Entropy (S) : entropy total dataset
- 3) S_1, S_2 : subset data hasil pemisahan atribut A
- 4) $|S_1|, |S_2|$: jumlah data pada masing-masing subset
- 5) $|S|$: jumlah total data
- 6) Info : Atribut dengan nilai Gain terbesar dipilih sebagai node pada *Decision Tree*

3.2.5. Pelatihan Model

Pelatihan model merupakan tahapan yang dilakukan untuk melatih data agar sesuai dengan kebutuhan proses pemodelan yang digunakan. Pada tahap ini, data yang telah dibersihkan dipersiapkan dan diatur sehingga dapat diproses secara optimal oleh metode yang diterapkan. Proses pelatihan dilakukan dengan mengelompokkan data berdasarkan setiap variabel yang digunakan sehingga pola hubungan antara variabel dan kategori kepuasan dapat dipelajari dengan baik oleh model. Adapun tabel yang digunakan pada tahap pelatihan model yaitu sebagai berikut.

Atribut	Partisi	Kasus (S)	Tidak Puas (S1)	Puas (S2)
Rasa	Enak	28	3	25
	Kurang Enak	22	20	2

Atribut	Partisi	Kasus (S)	Tidak Puas (S1)	Puas (S2)
Harga Menu	Murah	28	5	23
	Mahal	22	18	4

Atribut	Partisi	Kasus (S)	Tidak Puas (S1)	Puas (S2)
Pelayanan	Ramah	31	5	26
	Kurang Ramah	19	18	1

Atribut	Partisi	Kasus (S)	Tidak Puas (S1)	Puas (S2)
Kebersihan	Bersih	30	6	24
	Kurang Bersih	20	17	3

Atribut	Partisi	Kasus (S)	Tidak Puas (S1)	Puas (S2)
Kecepatan Penyajian	Cepat	27	2	25
	Lambat	23	21	2

Atribut	Partisi	Kasus (S)
Kategori	Puas	27
	Tidak Puas	23

Tabel di atas merupakan data yang telah dipisahkan berdasarkan setiap variabel yang digunakan, sehingga masing-masing atribut memiliki partisi data yang jelas dan terstruktur. Pemisahan ini dilakukan untuk melihat distribusi jumlah kasus, kategori puas, dan tidak puas pada setiap nilai atribut seperti rasa, harga menu, pelayanan, kebersihan, serta kecepatan penyajian. Dengan bentuk data seperti ini, hubungan antara setiap variabel dan tingkat kepuasan dapat diamati secara lebih rinci dan sistematis.

Data yang ditampilkan pada tabel tersebut merupakan data inti yang dibutuhkan dalam proses pengolahan dan perhitungan selanjutnya. Informasi jumlah kasus dan pembagian kategori kepuasan menjadi dasar dalam menghitung nilai entropy dan gain untuk setiap variabel. Oleh karena itu, data ini berperan penting dalam pembentukan model dan menjadi landasan utama dalam menentukan atribut yang paling berpengaruh terhadap tingkat kepuasan.

3.2.6. Klasifikasi Data

Klasifikasi data merupakan tahapan yang dilakukan untuk mengelompokkan data berdasarkan hasil perhitungan yang telah ditentukan. Pada tahap ini, proses perhitungan dilakukan menggunakan nilai entropy dan information gain untuk setiap variabel yang digunakan. Hasil dari perhitungan tersebut digunakan untuk mengetahui pola pembagian data ke dalam kategori puas dan tidak puas, sehingga dapat membantu menentukan keputusan klasifikasi secara sistematis dan terukur. Pada penelitian ini untuk proses perhitungan yang akan dilakukan dengan menggunakan metode *Decision Tree* yaitu sebagai berikut.

$$\text{Total data (S)} = 50$$

$$\text{Jumlah Kasus "Tidak Puas"} = 23$$

$$\text{Jumlah Kasus "Puas"} = 27$$

$$\begin{aligned} \text{Entropy (Total)} &= \left(-\frac{27}{50} \times \text{Log}_2 \left(\frac{27}{50} \right) \right) + \left(-\frac{23}{50} \times \text{Log}_2 \left(\frac{23}{50} \right) \right) \\ &= \left(-\frac{27}{50} \times \left(\frac{\ln(\frac{27}{50})}{\ln(2)} \right) \right) + \left(-\frac{23}{50} \times \left(\frac{\ln(\frac{23}{50})}{\ln(2)} \right) \right) \\ &= \left(-0,54 \times \left(\frac{-0,61}{0,69} \right) \right) + \left(-0,46 \times \left(\frac{-0,77}{0,69} \right) \right) \\ &= (-0,4) \times (-0,88) + (-0,46) \times (-1,11) \\ &= 0,47 + 0,51 \\ &= 0,98 \end{aligned}$$

Menghitung Entropy dan Gain Variabel Rasa

Menghitung Entropy Rasa (Enak)

$$\begin{aligned}
\text{Entropy (Total)} &= \left(-\frac{3}{28} \times \text{Log}_2\left(\frac{3}{28}\right)\right) + \left(-\frac{25}{28} \times \text{Log}_2\left(\frac{25}{28}\right)\right) \\
&= \left(-\frac{3}{28} \times \left(\frac{\ln\left(\frac{3}{28}\right)}{\ln(2)}\right)\right) + \left(-\frac{25}{28} \times \left(\frac{\ln\left(\frac{25}{28}\right)}{\ln(2)}\right)\right) \\
&= \left(-0,1 \times \left(\frac{-2,23}{0,69}\right)\right) + \left(-0,89 \times \left(\frac{-0,11}{0,69}\right)\right) \\
&= (-0,1) \times (-3,23) + (-0,89) \times (-0,15) \\
&= 0,32 + 0,13 \\
&= 0,45
\end{aligned}$$

Menghitung Entropy Rasa (Kurang Enak)

$$\begin{aligned}
\text{Entropy (Total)} &= \left(-\frac{20}{22} \times \text{Log}_2\left(\frac{20}{22}\right)\right) + \left(-\frac{2}{22} \times \text{Log}_2\left(\frac{2}{22}\right)\right) \\
&= \left(-\frac{20}{22} \times \left(\frac{\ln\left(\frac{20}{22}\right)}{\ln(2)}\right)\right) + \left(-\frac{2}{22} \times \left(\frac{\ln\left(\frac{2}{22}\right)}{\ln(2)}\right)\right) \\
&= \left(-0,9 \times \left(\frac{-0,09}{0,69}\right)\right) + \left(-0,09 \times \left(\frac{-2,39}{0,69}\right)\right) \\
&= (-0,9) \times (-0,13) + (-0,09) \times (-3,46) \\
&= 0,11 + 0,31 \\
&= 0,42
\end{aligned}$$

Menghitung Gain Rasa

$$\begin{aligned}
\text{Entropy (Total)} &= 0,98 - \left[\left(\frac{28}{50} \times (0,45)\right) + \left(\frac{22}{50} \times (0,42)\right)\right] \\
&= 0,98 - [(0,56 \times 0,45) + (0,44 \times 0,42)] \\
&= 0,98 - (0,25) + (0,18) \\
&= 0,98 - 0,43 \\
&= 0,55
\end{aligned}$$

Menghitung Entropy dan Gain Variabel Harga Menu

Menghitung Entropy Harga Menu (Murah)

$$\begin{aligned}\text{Entropy (Total)} &= \left(-\frac{5}{28} \times \text{Log}_2\left(\frac{5}{28}\right)\right) + \left(-\frac{23}{28} \times \text{Log}_2\left(\frac{23}{28}\right)\right) \\ &= \left(-\frac{5}{28} \times \left(\frac{\ln\left(\frac{5}{28}\right)}{\ln(2)}\right)\right) + \left(-\frac{23}{28} \times \left(\frac{\ln\left(\frac{23}{28}\right)}{\ln(2)}\right)\right) \\ &= \left(-0,17 \times \left(\frac{-1,72}{0,69}\right)\right) + \left(-0,82 \times \left(\frac{-0,19}{0,69}\right)\right) \\ &= (-0,17) \times (-2,49) + (-0,82) \times (-0,27) \\ &= 0,42 + 0,22 \\ &= 0,64\end{aligned}$$

Menghitung Entropy Harga Menu (Mahal)

$$\begin{aligned}\text{Entropy (Total)} &= \left(-\frac{18}{28} \times \text{Log}_2\left(\frac{18}{28}\right)\right) + \left(-\frac{4}{28} \times \text{Log}_2\left(\frac{4}{28}\right)\right) \\ &= \left(-\frac{18}{28} \times \left(\frac{\ln\left(\frac{18}{28}\right)}{\ln(2)}\right)\right) + \left(-\frac{4}{28} \times \left(\frac{\ln\left(\frac{4}{28}\right)}{\ln(2)}\right)\right) \\ &= \left(-0,64 \times \left(\frac{-0,44}{0,69}\right)\right) + \left(-0,4 \times \left(\frac{-1,94}{0,69}\right)\right) \\ &= (-0,64) \times (-0,63) + (-0,14) \times (-2,81) \\ &= 0,4 + 0,39 \\ &= 0,79\end{aligned}$$

Menghitung Gain Harga Menu

$$\begin{aligned}\text{Entropy (Total)} &= 0,98 - \left[\left(\frac{28}{50} \times 0,45\right) + \left(\frac{22}{50} \times 0,42\right)\right] \\ &= 0,98 - [(0,56 \times 0,64) + (0,44 \times 0,79)] \\ &= 0,98 - (0,35) + (0,34)\end{aligned}$$

$$= 0,98 - 0,69$$

$$= 0,29$$

Menghitung Entropy dan Gain Variabel Pelayanan

Menghitung Entropy Pelayanan (Ramah)

$$\begin{aligned} \text{Entropy (Total)} &= \left(-\frac{5}{31} \times \text{Log}_2\left(\frac{5}{31}\right)\right) + \left(-\frac{26}{31} \times \text{Log}_2\left(\frac{26}{31}\right)\right) \\ &= \left(-\frac{5}{31} \times \left(\frac{\ln\left(\frac{5}{31}\right)}{\ln(2)}\right)\right) + \left(-\frac{26}{31} \times \left(\frac{\ln\left(\frac{26}{31}\right)}{\ln(2)}\right)\right) \\ &= \left(-0,16 \times \left(\frac{-1,82}{0,69}\right)\right) + \left(-0,83 \times \left(\frac{-0,17}{0,69}\right)\right) \\ &= (-0,16) \times (-2,63) + (-0,83) \times (-0,24) \\ &= 0,42 + 0,19 \\ &= 0,61 \end{aligned}$$

Menghitung Entropy Pelayanan (Kurang Ramah)

$$\begin{aligned} \text{Entropy (Total)} &= \left(-\frac{18}{19} \times \text{Log}_2\left(\frac{18}{19}\right)\right) + \left(-\frac{1}{19} \times \text{Log}_2\left(\frac{1}{19}\right)\right) \\ &= \left(-\frac{18}{19} \times \left(\frac{\ln\left(\frac{18}{19}\right)}{\ln(2)}\right)\right) + \left(-\frac{1}{19} \times \left(\frac{\ln\left(\frac{1}{19}\right)}{\ln(2)}\right)\right) \\ &= \left(-0,94 \times \left(\frac{-0,05}{0,69}\right)\right) + \left(-0,05 \times \left(\frac{-2,94}{0,69}\right)\right) \\ &= (-0,64) \times (-0,63) + (-0,14) \times (-2,81) \\ &= 0,06 + 0,21 \\ &= 0,27 \end{aligned}$$

Menghitung Gain Pelayanan

$$\text{Entropy (Total)} = 0,98 - \left[\left(\frac{31}{50} \times 0,61\right) + \left(\frac{19}{50} \times 0,27\right)\right]$$

$$\begin{aligned}
&= 0,98 - [(0,62 \times 0,61) + (0,38 \times 0,25)] \\
&= 0,98 - (0,37) + (0,1) \\
&= 0,98 - 0,47 \\
&= 0,51
\end{aligned}$$

Menghitung Entropy dan Gain Variabel Kebersihan

Menghitung Entropy Kebersihan (Bersih)

$$\begin{aligned}
\text{Entropy (Total)} &= \left(-\frac{6}{30} \times \text{Log}2\left(\frac{6}{30}\right)\right) + \left(-\frac{24}{30} \times \text{Log}2\left(\frac{24}{30}\right)\right) \\
&= \left(-\frac{6}{30} \times \left(\frac{\ln\left(\frac{6}{30}\right)}{\ln(2)}\right)\right) + \left(-\frac{24}{30} \times \left(\frac{\ln\left(\frac{24}{30}\right)}{\ln(2)}\right)\right) \\
&= \left(-0,2 \times \left(\frac{-1,6}{0,69}\right)\right) + \left(-0,8 \times \left(\frac{-0,22}{0,69}\right)\right) \\
&= (-0,2) \times (-2,31) + (-0,8) \times (-0,31) \\
&= 0,46 + 0,24 \\
&= 0,7
\end{aligned}$$

Menghitung Entropy Kebersihan (Kurang Bersih)

$$\begin{aligned}
\text{Entropy (Total)} &= \left(-\frac{17}{20} \times \text{Log}2\left(\frac{17}{20}\right)\right) + \left(-\frac{3}{20} \times \text{Log}2\left(\frac{3}{20}\right)\right) \\
&= \left(-\frac{17}{20} \times \left(\frac{\ln\left(\frac{17}{20}\right)}{\ln(20)}\right)\right) + \left(-\frac{1}{20} \times \left(\frac{\ln\left(\frac{3}{20}\right)}{\ln(2)}\right)\right) \\
&= \left(-0,85 \times \left(\frac{-0,16}{0,69}\right)\right) + \left(-0,15 \times \left(\frac{-1,89}{0,69}\right)\right) \\
&= (-0,85) \times (-0,23) + (-0,15) \times (-2,73) \\
&= 0,19 + 0,4 \\
&= 0,59
\end{aligned}$$

Menghitung Gain Kebersihan

$$\begin{aligned}\text{Entropy (Total)} &= 0,98 - \left[\left(\frac{30}{50} \times 0,61 \right) + \left(\frac{20}{50} \times 0,27 \right) \right] \\ &= 0,98 - [(0,6 \times 0,7) + (0,4 \times 0,59)] \\ &= 0,98 - (0,42) + (0,23) \\ &= 0,98 - 0,65 \\ &= 0,33\end{aligned}$$

Menghitung Entropy dan Gain Variabel Kebersihan

Menghitung Entropy Kecepatan Penyajian (Cepat)

$$\begin{aligned}\text{Entropy (Total)} &= \left(-\frac{2}{27} \times \text{Log}_2 \left(\frac{2}{27} \right) \right) + \left(-\frac{25}{27} \times \text{Log}_2 \left(\frac{25}{27} \right) \right) \\ &= \left(-\frac{2}{27} \times \left(\frac{\ln\left(\frac{2}{27}\right)}{\ln(2)} \right) \right) + \left(-\frac{25}{27} \times \left(\frac{\ln\left(\frac{25}{27}\right)}{\ln(2)} \right) \right) \\ &= \left(-0,07 \times \left(\frac{-2,6}{0,69} \right) \right) + \left(-0,92 \times \left(\frac{-0,07}{0,69} \right) \right) \\ &= (-0,07) \times (-3,76) + (-0,92) \times (-0,1) \\ &= 0,26 + 0,09 \\ &= 0,35\end{aligned}$$

Menghitung Entropy Kecepatan Penyajian (Lambat)

$$\begin{aligned}\text{Entropy (Total)} &= \left(-\frac{21}{23} \times \text{Log}_2 \left(\frac{21}{23} \right) \right) + \left(-\frac{2}{23} \times \text{Log}_2 \left(\frac{2}{23} \right) \right) \\ &= \left(-\frac{21}{23} \times \left(\frac{\ln\left(\frac{21}{23}\right)}{\ln(2)} \right) \right) + \left(-\frac{2}{23} \times \left(\frac{\ln\left(\frac{2}{23}\right)}{\ln(2)} \right) \right) \\ &= \left(-0,91 \times \left(\frac{-0,05}{0,69} \right) \right) + \left(-0,08 \times \left(\frac{-2,44}{0,69} \right) \right) \\ &= (-0,91) \times (-0,13) + (-0,08) \times (-3,53)\end{aligned}$$

$$= 0,11 + 0,28$$

$$= 0,39$$

Menghitung Gain Kecepatan Penyajian

$$\text{Entropy (Total)} = 0,98 - \left[\left(\frac{27}{50} \times 0,35 \right) + \left(\frac{23}{50} \times 0,39 \right) \right]$$

$$= 0,98 - [(0,54 \times 0,35) + (0,46 \times 0,39)]$$

$$= 0,98 - (0,18) + (0,17)$$

$$= 0,98 - 0,35$$

$$= 0,63$$

Tulisan ini merupakan hasil perhitungan yang dilakukan dengan menggunakan metode *Decision Tree* untuk mengolah data tingkat kepuasan. Perhitungan dilakukan secara sistematis melalui tahapan pengukuran entropy dan information gain pada setiap variabel yang digunakan. Hasil dari proses tersebut kemudian diolah oleh komputer untuk menentukan pola dan aturan keputusan yang digunakan dalam proses klasifikasi, sehingga dapat menghasilkan model yang mampu memprediksi tingkat kepuasan secara terstruktur dan konsisten.

Tabel 3. 4. Hasil Perhitungan Entropy dan Gain

Atribut	Partisi	Kasus (S)	Tidak Puas (S1)	Puas (S2)	Entropy	Gain
Total		50	23	27	0.995	
Rasa						0.527
	Enak	28	3	25	0.491	
	Kurang Enak	22	20	2	0.439	
Harga Menu						0.315
	Murah	28	5	23	0.677	
	Mahal	22	18	4	0.684	
Pelayanan						0.487
	Ramah	31	5	26	0.637	
	Kurang Ramah	19	18	1	0.297	
Kebersihan						0.318

	Bersih	30	6	24	0.722	
	Kurang Bersih	20	17	3	0.610	
Kecepatan Penyajian						0.594
	Cepat	27	2	25	0.381	
	Lambat	23	21	2	0.426	

Tabel di atas menampilkan hasil perhitungan nilai entropy dan information gain dari seluruh data yang digunakan. Pada bagian total, terlihat bahwa jumlah data keseluruhan adalah 50 dengan distribusi 27 data puas dan 23 data tidak puas, sehingga menghasilkan nilai entropy sebesar 0,995. Nilai entropy ini menunjukkan tingkat ketidakpastian data sebelum dilakukan pemisahan berdasarkan atribut tertentu. Semakin mendekati nilai 1, maka data memiliki tingkat ketidakpastian yang cukup tinggi dan membutuhkan pemisahan lebih lanjut untuk memperoleh klasifikasi yang lebih jelas.

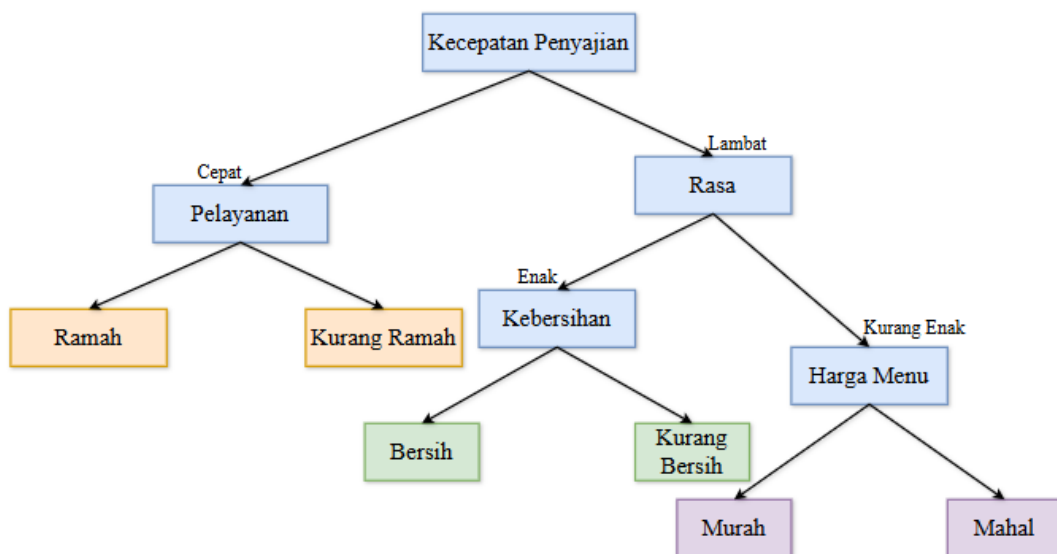
Setiap atribut pada tabel kemudian dihitung nilai entropy pada masing-masing partisinya, seperti Enak dan Kurang Enak pada atribut rasa, Murah dan Mahal pada harga menu, serta partisi lainnya. Nilai entropy pada setiap partisi menunjukkan seberapa homogen data dalam kategori tersebut. Misalnya, pada atribut rasa dengan partisi Enak, nilai entropy lebih rendah dibandingkan kondisi awal, yang menandakan bahwa data cenderung lebih terkelompok pada kategori puas. Perhitungan ini dilakukan untuk seluruh atribut guna melihat perubahan tingkat ketidakpastian setelah data dipisahkan.

Nilai information gain yang ditampilkan pada tabel merupakan selisih antara entropy total dengan entropy setelah pemisahan berdasarkan masing-masing atribut. Nilai gain yang lebih besar menunjukkan bahwa atribut tersebut memiliki

pengaruh yang lebih kuat dalam membedakan kategori puas dan tidak puas. Berdasarkan tabel, kecepatan penyajian memiliki nilai gain tertinggi, diikuti oleh rasa dan pelayanan, sehingga atribut-atribut tersebut dinilai paling berperan dalam proses klasifikasi tingkat kepuasan. Hasil ini menjadi dasar dalam pembentukan pohon keputusan pada metode *Decision Tree*.

3.2.7. Analisis Hasil

Analisis hasil dilakukan untuk menafsirkan keluaran dari proses perhitungan dan pemodelan yang telah dilakukan. Selain melihat nilai entropy dan information gain, hasil lainnya yang diperoleh berupa pohon keputusan yang menggambarkan alur pengambilan keputusan berdasarkan atribut-atribut yang digunakan. Pohon keputusan tersebut menunjukkan urutan atribut yang paling berpengaruh dalam menentukan tingkat kepuasan, serta aturan-aturan klasifikasi yang terbentuk dari data. Dengan adanya pohon keputusan, hasil analisis menjadi lebih mudah dipahami karena hubungan antar variabel dan kategori kepuasan dapat divisualisasikan secara jelas dan sistematis.



Gambar 3. 1. Hasil Pohon Keputusan

Pohon keputusan dibangun berdasarkan urutan nilai information gain pada tabel, di mana atribut dengan nilai gain tertinggi ditempatkan sebagai node awal. Berdasarkan tabel, kecepatan penyajian memiliki nilai gain paling besar yaitu 0,594, sehingga atribut ini menjadi akar pohon keputusan. Pembagian data ke dalam kategori cepat dan lambat mampu menurunkan nilai entropy secara signifikan, karena pada partisi cepat mayoritas data berada pada kategori puas, sedangkan pada partisi lambat mayoritas data berada pada kategori tidak puas. Hal inilah yang menjadikan kecepatan penyajian sebagai faktor utama dalam menentukan tingkat kepuasan.

Pada cabang kecepatan penyajian cepat, atribut pelayanan ditempatkan pada tingkat berikutnya. Pemilihan pelayanan sebagai node lanjutan didasarkan pada nilai gain pelayanan yang cukup tinggi yaitu 0,487. Pada kondisi penyajian cepat, pelayanan yang ramah cenderung menghasilkan kepuasan yang lebih tinggi, sementara pelayanan yang kurang ramah dapat menurunkan tingkat kepuasan. Struktur ini menunjukkan bahwa meskipun kecepatan sudah baik, sikap pelayanan tetap berperan penting dalam membentuk pengalaman pengunjung.

Sementara itu, pada cabang kecepatan penyajian lambat, atribut rasa digunakan sebagai node berikutnya karena memiliki nilai gain yang tinggi yaitu 0,527. Jika rasa dinilai enak, maka atribut kebersihan digunakan sebagai pemisah lanjutan, sedangkan jika rasa kurang enak, atribut harga menu menjadi faktor pembeda berikutnya. Penempatan kebersihan dan harga menu pada tingkat bawah pohon keputusan sesuai dengan nilai gain keduanya yang lebih rendah dibandingkan atribut lain. Dengan demikian, struktur pohon keputusan sepenuhnya

mencerminkan urutan dan besarnya nilai gain pada tabel, sehingga setiap keputusan yang dihasilkan memiliki dasar perhitungan yang jelas dan logis.

3.2.8. Evaluasi Model

Evaluasi metode Decision Tree pada penelitian ini menggunakan Confusion Matrix. Hasil tersebut yang nantinya akan memberikan nilai efektivitas dari metode yang digunakan.

		Predicted		Σ
		Puas	Tidak Puas	
Actual	Puas	27	0	27
	Tidak Puas	0	23	23
Σ		27	23	50

Hasil True Positive (TP) adalah 27. True Negative (TN) adalah 23, False Positive (FP) adalah 0 dan False Negative (FN) adalah 0. Maka Nilai akurasi, presisi dan recall adalah sebagai berikut:

$$\text{Accuracy} = \frac{27+23}{27+23+0+0} \times 100\% \quad \text{Then the Accuracy value} = 100\%$$

$$\text{Presisi} = \frac{27}{27+0} \times 100\% \quad \text{Then the Precision value} = 100\%$$

$$\text{Recall} = \frac{27}{27+0} \times 100\% \quad \text{Then the Recall value} = 100\%$$

$$\text{F1 - Score} = 2 \times \frac{100 \times 100}{100+100} = 2 \times \frac{10000}{200} = 2 \times \frac{10000}{200} = 2 \times 50 = 100\%$$

Hasil evaluasi menggunakan Confusion Matrix menunjukkan bahwa metode Decision Tree mampu melakukan klasifikasi tingkat kepuasan secara sempurna. Nilai True Positive sebesar 27 dan True Negative sebesar 23 menandakan bahwa seluruh data berhasil diklasifikasikan dengan benar tanpa adanya kesalahan prediksi, karena nilai False Positive dan False Negative sama-sama bernilai 0.

Kondisi tersebut menghasilkan nilai akurasi sebesar 100%, yang berarti seluruh data teridentifikasi sesuai kategorinya. Nilai presisi dan recall juga mencapai 100%, yang menunjukkan bahwa model mampu mengenali seluruh data positif tanpa kesalahan dan mampu mendeteksi seluruh kasus positif dengan tepat. Nilai F1-Score yang mencapai 100% menguatkan bahwa model ini memiliki keseimbangan performa antara presisi dan recall, sehingga dapat disimpulkan bahwa metode Decision Tree bekerja dengan sangat efektif untuk memprediksi tingkat kepuasan pada data yang digunakan.