

BAB II

LANDASAN TEORI

2.1. *Knowledge Discovery in Database (KDD)*

Knowledge Discovery in Database (KDD) merupakan proses sistematis untuk menemukan pola, informasi penting, dan pengetahuan baru dari kumpulan data yang besar (Mauliadi, 2022). Tahapan KDD meliputi seleksi data, pembersihan data, transformasi, proses data mining, serta interpretasi hasil. Setiap tahap memiliki peran penting dalam memastikan bahwa pengetahuan yang dihasilkan benar-benar valid, akurat, dan dapat digunakan untuk pengambilan keputusan. KDD biasanya memanfaatkan teknik data mining seperti klasifikasi, klustering, atau asosiasi untuk menemukan pola tersembunyi yang tidak mudah dilihat secara manual (Suarna et al., 2024). Dengan adanya perangkat lunak seperti RapidMiner, proses KDD menjadi lebih mudah karena setiap tahap dapat dilakukan secara terstruktur dan terotomatisasi melalui alur kerja (*workflow*) yang jelas, sehingga pengguna dapat mengolah data dalam jumlah besar dengan lebih efisien dan terukur (Nugraha et al., 2023).

Penerapan KDD memberikan manfaat besar bagi upaya mengungkap pola kepuasan pengunjung di *Cafe Teman Nongkrong*, terutama ketika data yang digunakan bersifat beragam dan kompleks. Melalui RapidMiner, setiap tahapan mulai dari pemilihan data survei pelanggan, pembersihan nilai yang tidak lengkap, hingga proses klasifikasi menggunakan metode *Decision Tree* dapat dilakukan secara terpadu (Maulana et al., 2023). Hasil data mining yang dihasilkan dari proses ini dapat membantu mengidentifikasi variabel mana yang paling memengaruhi

tingkat kepuasan pengunjung, sehingga pengelola kafe dapat memperoleh gambaran yang lebih jelas mengenai perilaku pelanggan (Ameliana et al., 2024).

2.1.1. Konsep Dasar *Knowledge Discovery in Database (KDD)*

Knowledge Discovery in Database (KDD) merupakan proses komprehensif untuk menemukan pengetahuan baru dari kumpulan data yang besar melalui serangkaian tahapan yang terstruktur (Fajar et al., 2024). Proses ini meliputi pemilihan data, pembersihan, transformasi, penerapan algoritma data mining, hingga interpretasi hasil. KDD bertujuan mengubah data mentah menjadi informasi bermakna yang dapat digunakan sebagai dasar pengambilan keputusan (Sitanggang et al., 2024). Setiap tahap memiliki peran penting, mulai dari memastikan kualitas data, memilih metode analisis yang tepat, hingga menafsirkan pola atau hubungan yang ditemukan. Dengan kemampuannya mengekstraksi pola tersembunyi, KDD menjadi fondasi dari berbagai pendekatan analisis modern, terutama ketika data berjumlah besar, kompleks, dan tidak mudah dianalisis secara manual (Ubaid et al., 2025).

Penerapan KDD membantu menguraikan pola kepuasan pengunjung di *Cafe Teman Nongkrong* melalui tahapan mulai dari pengumpulan data survei, pembersihan data menggunakan teknik *Preprocessing*, hingga proses analisis menggunakan algoritma seperti *Decision Tree* yang dijalankan melalui perangkat lunak RapidMiner (Julianti et al., 2024). Setiap tahap KDD berperan penting dalam menghasilkan informasi yang akurat, misalnya memastikan data penilaian pelanggan bebas dari duplikasi, menyiapkan atribut yang relevan, serta mengolahnya menjadi model prediksi yang dapat mengelompokkan pengunjung

berdasarkan tingkat kepuasan (Atalya et al., 2024). Dengan memanfaatkan alur kerja KDD di RapidMiner, pengelola kafe dapat memperoleh wawasan yang lebih mendalam tentang faktor yang paling berpengaruh terhadap pengalaman pelanggan, sehingga strategi peningkatan layanan dapat dirumuskan secara lebih terarah dan berbasis bukti.

2.2. Teori dan Konsep Metode *Decision Tree* untuk Prediksi Tingkat Kepuasan

Teori mengenai kepuasan pelanggan menjelaskan bahwa kepuasan terbentuk dari perbandingan antara harapan awal dan pengalaman nyata yang diterima oleh pelanggan. Jika pengalaman sesuai atau melebihi harapan, maka kepuasan cenderung meningkat. Faktor-faktor seperti kualitas produk, pelayanan, suasana lingkungan, dan persepsi harga menjadi elemen utama yang memengaruhi terbentuknya kepuasan tersebut. Selain itu, pendekatan analisis data dan *Machine learning* memberikan dasar teoritis bahwa pola perilaku pelanggan dapat dimodelkan dan diprediksi melalui algoritma yang mampu mengenali hubungan antar variabel secara sistematis. Salah satu metode yang sering digunakan adalah *Decision Tree*, yang menyajikan alur keputusan dalam bentuk struktur pohon sehingga memudahkan interpretasi hubungan sebab-akibat antar faktor.

Hubungan antara teori tersebut dengan pembahasan terletak pada kebutuhan untuk memahami variabel-variabel yang memengaruhi pengalaman pengunjung di *Cafe Teman Nongkrong*. Dengan menggabungkan konsep kepuasan pelanggan dan kemampuan prediktif *Decision Tree*, analisis dapat dilakukan untuk mengidentifikasi faktor dominan yang menentukan kepuasan pengunjung. Pendekatan ini memungkinkan pemetaan pola kepuasan secara lebih objektif

melalui data yang diperoleh dari penilaian dan pengalaman pelanggan, sehingga hasil yang dihasilkan dapat mendukung pengelola dalam menentukan strategi peningkatan layanan yang lebih tepat dan efektif.

2.3. Teori Tingkat Kepuasan

Tingkat kepuasan merupakan hasil evaluasi pengunjung terhadap pengalaman yang mereka terima setelah menikmati suatu layanan (Andi et al., 2022). Kepuasan muncul ketika kualitas produk, pelayanan, harga, dan suasana mampu memenuhi atau melampaui harapan yang telah terbentuk sebelumnya. Faktor-faktor seperti rasa makanan dan minuman, kecepatan penyajian, kebersihan, kenyamanan ruang, serta interaksi dengan staf menjadi penentu utama dalam pembentukan kepuasan (Arisoemaryo et al., 2022). Ketidaksesuaian di salah satu aspek saja dapat memengaruhi penilaian keseluruhan pengunjung, sehingga pemahaman terhadap elemen penyusun kepuasan menjadi hal yang penting dalam pengelolaan layanan (Sulistyo et al., 2022).

Penerapan konsep kepuasan sangat relevan dengan kondisi yang terjadi di *Cafe Teman Nongkrong*, di mana pengunjung menunjukkan variasi pengalaman terhadap kualitas produk dan layanan. Perbedaan persepsi terkait rasa menu, harga, maupun konsistensi pelayanan membuat tingkat kepuasan tidak seragam. Untuk itu, pemahaman menyeluruh tentang faktor pembentuk kepuasan diperlukan agar dapat diidentifikasi variabel yang paling berpengaruh (Pandemi, 2022). Melalui analisis yang lebih terstruktur menggunakan metode seperti *Decision Tree*, pola kepuasan dapat dipetakan sehingga pengelola memiliki dasar yang lebih tepat dalam

meningkatkan mutu layanan dan menciptakan pengalaman yang lebih memuaskan bagi pengunjung (Mahendro et al., 2022).

2.2.1. Persepsi Terhadap Kualitas Produk

Persepsi terhadap kualitas produk terbentuk dari berbagai aspek yang dirasakan langsung oleh pelanggan, tidak hanya terbatas pada rasa makanan dan minuman, tetapi juga mencakup harga, kebersihan, serta penyajian secara keseluruhan. Kualitas rasa menjadi faktor utama karena cita rasa yang enak, konsisten, dan sesuai selera akan meninggalkan kesan positif bagi pelanggan. Selain itu, kesesuaian harga dengan kualitas yang diterima turut memengaruhi penilaian, di mana pelanggan akan merasa puas apabila harga yang dibayarkan sebanding dengan rasa, porsi, dan pengalaman yang diperoleh. Kebersihan juga menjadi bagian penting dari kualitas produk, baik kebersihan makanan, peralatan makan, maupun area penyajian, karena hal tersebut berkaitan langsung dengan kenyamanan dan kepercayaan pelanggan. Ketika sebuah kafe mampu menjaga kualitas rasa, menetapkan harga yang wajar, serta mempertahankan standar kebersihan dan penyajian yang baik, pelanggan akan menilai produk yang ditawarkan memiliki kualitas tinggi dan pada akhirnya mendorong terbentuknya kepuasan serta loyalitas.

Persepsi kualitas produk memiliki hubungan langsung dengan tingkat kepuasan pengunjung di *Cafe Teman Nongkrong*. Rasa makanan dan minuman yang dianggap enak sering menjadi alasan utama sebagian pengunjung merasa puas dan memberikan ulasan positif. Sebaliknya, apabila kualitas produk tidak stabil atau kurang sesuai dengan harapan, maka muncul peluang terjadinya ketidakpuasan

meskipun aspek lain sudah mendukung. Faktor ini menjadi salah satu variabel penting yang diolah dan dianalisis dalam model *Decision Tree* untuk mengetahui seberapa besar pengaruh kualitas produk terhadap prediksi tingkat kepuasan pengunjung

2.2.2. Kualitas Pelayanan dan Interaksi dengan Staff

Kualitas pelayanan dan interaksi dengan staf menjadi elemen penting dalam membentuk pengalaman pengunjung di sebuah kafe. Pelayanan yang cepat, ramah, konsisten, dan penuh perhatian dapat menciptakan kesan positif yang mendorong kenyamanan, kecepatan penyajian pelayanan serta rasa dihargai. Sebaliknya, sikap staf yang kurang responsif, penyajian yang lambat, atau pelayanan yang tidak konsisten dapat menurunkan kualitas pengalaman meskipun aspek lain seperti rasa makanan sudah baik.

Pada Cafe Teman Nongkrong, kualitas pelayanan dan interaksi staf menjadi salah satu variabel yang turut memengaruhi tingkat kepuasan pengunjung. Beberapa pelanggan memberikan apresiasi terhadap keramahan staf, namun sebagian lainnya mengeluhkan ketidakkonsistenan pelayanan, seperti keterlambatan penyajian atau kurangnya perhatian pada saat ramai. Perbedaan pengalaman ini memunculkan variasi tingkat kepuasan yang perlu dianalisis lebih lanjut. Dengan memasukkan variabel pelayanan ke dalam model *Decision Tree*, dapat diketahui seberapa besar pengaruh interaksi staf terhadap kepuasan pelanggan serta bagaimana faktor ini berperan dalam membentuk pola prediksi.

2.4. Cafe Teman Nongkrong

Cafe Teman Nongkrong merupakan salah satu kafe yang hadir dengan konsep modern dan santai yang banyak diPuasi oleh masyarakat, terutama kalangan muda. Kafe ini menawarkan suasana nyaman dengan desain interior yang menarik, sehingga cocok digunakan untuk bersantai, mengobrol, hingga bekerja menggunakan perangkat pribadi. Menu yang disediakan cukup beragam, mulai dari makanan ringan, hidangan utama, hingga berbagai pilihan minuman yang menjadi daya tarik utama bagi pengunjung. Selain itu, pelayanan yang diberikan berupaya dibuat ramah dan responsif agar pengunjung merasa dihargai dan mendapatkan pengalaman yang menyenangkan selama berada di kafe. Kombinasi antara suasana, menu, dan layanan inilah yang membuat *Cafe Teman Nongkrong* menjadi salah satu pilihan favorit bagi pelanggan yang mencari tempat dengan suasana yang hangat dan nyaman.

Keberadaan *Cafe Teman Nongkrong* menjadi semakin menarik ketika dikaitkan dengan kebutuhan untuk memahami tingkat kepuasan pengunjung secara lebih mendalam. Setiap pengunjung yang datang memiliki ekspektasi yang berbeda, sehingga pengalaman mereka bisa beragam ada yang memuji kualitas rasa dan kenyamanan ruang, sementara sebagian lain menyoroti harga, pelayanan, atau variasi menu. Variasi pengalaman ini menuntut adanya pendekatan analitis berbasis data agar pengelola dapat mengetahui faktor mana saja yang paling memengaruhi kepuasan pengunjung. Melalui penerapan metode *Decision Tree*, pola kepuasan pelanggan dapat diidentifikasi secara lebih jelas, sehingga strategi peningkatan kualitas layanan dapat dirancang secara tepat, akurat, dan sesuai dengan kebutuhan serta karakteristik pengunjung *Cafe Teman Nongkrong*.

2.5. Metode yang Digunakan

Metode yang digunakan berfokus pada penerapan *Decision Tree* sebagai salah satu teknik klasifikasi dalam *Machine learning* yang mampu memetakan pola hubungan antar variabel secara terstruktur (Adipradana, 2023). Metode ini bekerja dengan membagi data ke dalam cabang-cabang keputusan berdasarkan nilai atribut yang paling memberikan informasi. Setiap cabang mewakili aturan yang terbentuk dari data, sementara daun pohon menjadi hasil klasifikasi akhir (Wibina et al., 2022). Keunggulan *Decision Tree* terletak pada kemudahannya untuk dipahami karena menghasilkan model yang bersifat visual dan logis (Cahyono et al., 2024). Selain itu, metode ini mampu menangani data dengan berbagai jenis atribut, baik kategorikal maupun numerik, serta dapat dievaluasi performanya menggunakan metrik seperti akurasi, presisi, dan recall untuk melihat sejauh mana model mampu mengenali pola dengan tepat (Wibowo et al., 2024).

Penerapan metode *Decision Tree* digunakan untuk mengolah data penilaian pengunjung *Cafe Teman Nongkrong* agar dapat memprediksi tingkat kepuasan berdasarkan variabel-variabel seperti rasa makanan, pelayanan, suasana tempat, harga, dan faktor-faktor lain yang relevan (Prasetyo et al., 2025). Melalui struktur pohon keputusan, metode ini membantu mengungkap variabel mana yang paling dominan memengaruhi kepuasan maupun ketidakpuasan pengunjung (Fihir et al., 2023). Model yang dihasilkan kemudian digunakan untuk mengelompokkan data ke dalam kategori tertentu sehingga pola kepuasan dapat diamati secara lebih jelas (Pramesti & Pratiwi, 2023). Dengan adanya model prediktif ini, proses identifikasi faktor utama menjadi lebih terarah dan berbasis data, sehingga pengelola dapat memperoleh gambaran yang lebih akurat mengenai perilaku pengunjung serta

merumuskan strategi peningkatan layanan dan kualitas kafe secara lebih efektif (Mediana & Yamasari, 2024).

2.4.1. Prinsip Dasar Metode *Decision Tree*

Metode *Decision Tree* bekerja berdasarkan prinsip pemecahan masalah secara bertahap melalui struktur pohon keputusan yang terdiri dari simpul akar, simpul cabang, dan simpul daun (Azwanti & Putria, 2024). Setiap simpul merepresentasikan suatu pengujian terhadap atribut tertentu, sedangkan cabang menggambarkan hasil dari pengujian tersebut, dan simpul daun menunjukkan keputusan atau klasifikasi akhir. Proses pembentukan pohon dimulai dari pemilihan atribut paling informatif menggunakan ukuran seperti *information gain* atau *gain ratio*. Atribut yang memberikan kemampuan terbaik dalam membagi data ke dalam kelas yang lebih homogen dipilih sebagai simpul akar. Selanjutnya, proses ini berlanjut secara rekursif hingga seluruh data terklasifikasi dengan baik atau hingga tidak ada lagi atribut yang dapat diproses. Sifatnya yang mudah dipahami, interpretatif, serta mampu menangani data kategorikal maupun numerik membuat *Decision Tree* banyak digunakan dalam analisis prediksi dan klasifikasi.

Penerapan prinsip *Decision Tree* sangat relevan untuk memetakan pola kepuasan pengunjung di *Cafe Teman Nongkrong* karena setiap faktor seperti rasa makanan, harga, pelayanan, kebersihan, dan kenyamanan dapat dianalisis sebagai atribut yang berperan dalam membentuk kepuasan. Melalui proses pemilihan atribut yang paling berpengaruh, pohon keputusan dapat menunjukkan variabel mana yang menjadi penentu utama perbedaan antara pengunjung yang merasa puas dan yang tidak puas. Struktur pohon yang dihasilkan mampu memberikan

gambaran visual tentang alur keputusan, mulai dari faktor paling dominan hingga kombinasi kondisi yang memengaruhi hasil klasifikasi (Luh et al., 2024).

2.4.2. Kelebihan Metode *Decision Tree*

Metode *Decision Tree* memiliki kelebihan karena mampu menyajikan hasil analisis dalam bentuk struktur pohon yang intuitif dan mudah dipahami. Setiap percabangan merepresentasikan keputusan berdasarkan nilai atribut tertentu, sehingga alur penentuan prediksi dapat diikuti secara jelas tanpa memerlukan pemahaman algoritmik yang kompleks. Kemampuan metode ini untuk menangani data kategorikal maupun numerik membuatnya fleksibel digunakan dalam berbagai jenis analisis. Selain itu, *Decision Tree* mampu mengidentifikasi atribut yang paling berpengaruh dalam pengambilan keputusan, sehingga menghasilkan model yang tidak hanya akurat, tetapi juga informatif. Sifatnya yang interpretatif menjadikannya salah satu metode *Machine learning* yang unggul ketika tujuan utama analisis adalah pemahaman pola, bukan sekadar akurasi prediksi (Bagas et al., n.d.).

Penerapan metode *Decision Tree* sangat relevan untuk menganalisis faktor-faktor yang memengaruhi tingkat kepuasan pengunjung di *Cafe Teman Nongkrong*. Struktur pohon keputusan memungkinkan pengelola melihat dengan jelas variabel mana yang paling dominan dalam memengaruhi kepuasan, seperti harga, pelayanan, atau kualitas menu (Sinaga et al., 2022). Dari percabangan yang terbentuk, pengelola dapat memahami urutan prioritas faktor yang memengaruhi kepuasan pengunjung sehingga langkah perbaikan dapat diarahkan dengan lebih tepat. Keunggulan metode ini dalam menampilkan pola pengaruh antar variabel

memungkinkan hasil analisis digunakan sebagai dasar pengambilan keputusan strategis untuk meningkatkan mutu layanan, kenyamanan, serta pengalaman keseluruhan pengunjung secara lebih efektif.

2.4.3. Kekurangan Metode *Decision Tree*

Metode *Decision Tree* memiliki beberapa kelemahan yang perlu diperhatikan dalam proses analisis. Struktur pohon keputusan cenderung menjadi terlalu kompleks apabila jumlah variabel dan data yang digunakan sangat banyak, sehingga menghasilkan model yang *overfitting* atau terlalu menyesuaikan diri dengan data latih. Kondisi ini dapat menyebabkan penurunan akurasi ketika model digunakan pada data baru. Selain itu, *Decision Tree* sangat sensitif terhadap perubahan kecil pada data; sedikit perbedaan pada nilai atribut dapat menghasilkan struktur pohon yang berbeda secara signifikan.

Kelemahan tersebut dapat memengaruhi proses analisis kepuasan pengunjung di *Cafe Teman Nongkrong*, terutama ketika data survei memiliki variasi jawaban yang beragam atau terdapat ketidakseimbangan antara kategori puas dan tidak puas. Struktur pohon yang terlalu dalam dapat menyebabkan hasil klasifikasi menjadi kurang stabil dan sulit diinterpretasikan. Sensitivitas terhadap perubahan data juga membuat model perlu diuji dan divalidasi dengan hati-hati agar prediksi yang dihasilkan tetap akurat. Meskipun begitu, metode ini tetap relevan digunakan karena tingkat interpretabilitasnya yang tinggi, namun perlu didukung dengan teknik pengaturan seperti *pruning* atau penggunaan data yang cukup representatif agar hasil analisis mengenai faktor-faktor kepuasan pengunjung lebih dapat diandalkan.

2.6. *Machine learning*

Machine learning merupakan cabang kecerdasan buatan yang memungkinkan sistem komputer belajar dari data dan melakukan prediksi atau pengambilan keputusan tanpa harus diprogram secara detail untuk setiap tugas (Zuama et al., 2022). Proses belajar ini dilakukan melalui identifikasi pola, keteraturan, dan hubungan antar variabel dalam data (Astuti & Komputer, 2024). Algoritma *Machine learning* bekerja dengan membangun model matematis berdasarkan data latih, kemudian menggunakannya untuk memprediksi hasil pada data baru. Metode ini memiliki kemampuan untuk meningkatkan akurasi seiring bertambahnya jumlah data yang dianalisis (Panji Rachmat Setiawan et al., 2022). Keunggulan *Machine learning* terletak pada kemampuannya menangani data yang kompleks, mendeteksi pola yang tidak terlihat oleh analisis manual, serta melakukan pemrosesan dalam skala besar dengan cepat dan efisien (Khoiri & Wahid, 2024).

Penerapan *Machine learning* sangat relevan untuk menganalisis tingkat kepuasan di *Cafe Teman Nongkrong*, karena data penilaian pengunjung biasanya memiliki variasi yang cukup tinggi dan melibatkan banyak faktor (Akbar et al., 2024). Dengan memanfaatkan algoritma seperti *Decision Tree*, pola hubungan antar variabel seperti rasa, harga, pelayanan, dan kenyamanan dapat dipetakan secara jelas sehingga menghasilkan model prediksi yang mudah dipahami (Alfarizi et al., 2023). Pendekatan ini memungkinkan pengelola kafe mendapatkan gambaran yang lebih akurat mengenai faktor dominan yang memengaruhi kepuasan, serta

membantu dalam merumuskan strategi peningkatan layanan yang lebih tepat sasaran. Penggunaan *Machine learning* memberikan kelebihan berupa hasil yang objektif, terukur, dan dapat digunakan sebagai dasar pengambilan keputusan berbasis data (Taufik et al., 2024).

2.6.1. Kosep Dasar *Machine Learning*

Machine learning merupakan cabang kecerdasan buatan yang memungkinkan sistem komputer belajar dari data dan meningkatkan performanya tanpa harus diprogram secara manual untuk setiap tugas (Sibuea & Widodo, 2024). Proses pembelajarannya dilakukan melalui identifikasi pola, hubungan antarvariabel, serta kecenderungan tertentu yang tersimpan di dalam data (Lofandri, 2023). Sistem kemudian membentuk model yang mampu membuat keputusan atau prediksi berdasarkan contoh-contoh yang telah dipelajari sebelumnya. Metode ini sangat bermanfaat ketika data yang dianalisis memiliki kompleksitas tinggi, jumlahnya besar, atau terdapat interaksi variabel yang sulit dijelaskan secara intuitif oleh manusia (Fahmuddin et al., 2023). Dengan kemampuan untuk bekerja secara adaptif, *Machine learning* menjadi alat analitis yang semakin banyak digunakan dalam berbagai bidang seperti bisnis, kesehatan, keuangan, hingga analisis perilaku konsumen (Agustina & Rahmah, 2022).

Penerapan *Machine learning* dapat membantu menggali pola kepuasan pengunjung di *Cafe Teman Nongkrong* dengan lebih akurat dan objektif (Mazyra et al., 2024). Data penilaian pelanggan yang beragam—mulai dari rasa, harga, pelayanan, hingga kenyamanan—dapat diproses untuk menemukan faktor yang paling berpengaruh terhadap tingkat kepuasan. Metode seperti *Decision Tree*

memungkinkan model mengklasifikasikan tingkat kepuasan berdasarkan hubungan antar variabel yang ditemukan selama proses pembelajaran (Syahputra & Hanifah, n.d.). Kemampuan *Machine learning* untuk memetakan pola dari data aktual memberikan peluang bagi pengelola kafe untuk memahami preferensi pengunjung dengan lebih mendalam, sehingga strategi peningkatan layanan dapat dirancang secara lebih tepat dan efektif (Atmojo et al., 2024).

2.6.2. Langkah-Langkah dalam *Machine Learning*

Kalimat di bawah ini merupakan rangkaian langkah yang digunakan dalam proses *Machine learning* untuk membangun model prediktif secara sistematis. Tahapan berikut menjelaskan alur kerja mulai dari pengumpulan data hingga evaluasi model yang digunakan.

1. Tahap Pengumpulan Data

Tahap pengumpulan data merupakan langkah awal yang sangat penting karena kualitas data menentukan kualitas model yang akan dibangun. Pada tahapan ini, seluruh data yang dibutuhkan dikumpulkan dari sumber yang relevan, seperti hasil distribusi kuesioner berupa google form. Data yang diperoleh biasanya mencakup berbagai variabel seperti penilaian rasa, harga, pelayanan, suasana, maupun variabel lain yang berhubungan dengan kepuasan. Semakin lengkap dan representatif data yang dikumpulkan, semakin besar peluang model *Machine learning* menghasilkan prediksi yang akurat. Oleh karena itu, proses pengumpulan harus dilakukan secara terstruktur dan memastikan bahwa setiap data yang diambil benar-benar sesuai dengan tujuan analisis.

2. Tahap *Preprocessing Data*

Tahap *Preprocessing Data* dilakukan untuk membersihkan serta mempersiapkan data sebelum masuk ke proses pemodelan. Data mentah biasanya mengandung nilai kosong, data duplikasi, atau format yang tidak konsisten sehingga perlu dilakukan pembersihan untuk menjaga kualitas data. Selain itu, pada tahap ini dapat dilakukan proses transformasi seperti normalisasi, standarisasi, pengkodean data kategorikal, atau penghapusan outlier agar data lebih siap diproses oleh algoritma *Machine learning*. *Preprocessing* yang baik memastikan bahwa model tidak terganggu oleh noise atau kesalahan data, sehingga hasil klasifikasi atau prediksi menjadi lebih akurat dan dapat diandalkan.

3. Tahap Pembagian Data

Setelah data melewati tahap *Preprocessing*, langkah berikutnya adalah membagi data menjadi dua bagian utama, yaitu data pelatihan (training set) dan data pengujian (testing set). Pembagian ini bertujuan untuk memastikan bahwa model dapat belajar dari sebagian data, kemudian diuji menggunakan data lain yang belum pernah dilihat sebelumnya. Dengan cara ini, kemampuan model dalam melakukan generalisasi dapat dievaluasi secara objektif. Proporsi pembagian biasanya berkisar antara 70:30 atau 80:20 tergantung kebutuhan analisis. Tahapan ini sangat penting untuk mencegah masalah seperti overfitting, yaitu kondisi ketika model terlalu menyesuaikan diri dengan data pelatihan sehingga performanya menurun saat digunakan pada data baru.

4. Tahap Perancangan Model

Tahap perancangan model melibatkan pemilihan metode *Machine learning* yang akan digunakan, seperti *Decision Tree*, serta proses membangun model

berdasarkan data pelatihan. Pada tahap ini, algoritma menganalisis variabel-variabel yang tersedia untuk menemukan pola, aturan keputusan, dan hubungan antar faktor yang berpengaruh terhadap variabel target. Struktur model kemudian dibentuk sesuai dengan prinsip kerja algoritma, misalnya pembentukan node dan cabang pada *Decision Tree* yang menggambarkan proses pengambilan keputusan. Tahapan ini menentukan seberapa baik model dapat mempelajari karakteristik data sehingga mampu menghasilkan prediksi yang sesuai dengan pola yang ada.

5. Tahap Evaluasi Model

Tahap evaluasi model bertujuan untuk mengukur seberapa baik performa model setelah dibangun. Proses ini dilakukan dengan menguji model menggunakan data pengujian dan menghitung metrik evaluasi seperti akurasi, presisi, recall, atau F1-score. Evaluasi diperlukan untuk memastikan bahwa model tidak hanya bekerja baik pada data pelatihan tetapi juga efektif menghadapi data baru. Melalui evaluasi yang menyeluruh, dapat diketahui apakah model sudah cukup layak digunakan atau perlu dilakukan perbaikan seperti penyesuaian parameter, penggunaan fitur tambahan, atau perubahan metode. Evaluasi model merupakan langkah krusial untuk memastikan hasil prediksi yang dihasilkan benar-benar dapat dipercaya.

2.7. Model Klasifikasi

Model klasifikasi merupakan salah satu pendekatan dalam *Machine learning* yang digunakan untuk mengelompokkan data ke dalam kelas atau kategori tertentu berdasarkan pola yang ditemukan dari data sebelumnya. Proses klasifikasi melibatkan pembelajaran dari sekumpulan data berlabel sehingga algoritma dapat mengenali hubungan antara fitur-fitur masukan dengan kategori yang menjadi

target. Ketika model mendapatkan data baru, model akan menentukan kelas yang paling sesuai berdasarkan pola yang telah dipelajari. Teknik klasifikasi banyak digunakan karena mampu menangani data yang kompleks, memiliki kemampuan prediksi yang kuat, serta dapat memberikan pemahaman mengenai faktor-faktor yang memengaruhi suatu keputusan. Berbagai algoritma seperti *Decision Tree*, *Random Forest*, *Naïve Bayes*, dan *Support Vector Machine* sering digunakan dalam proses klasifikasi dengan kelebihan masing-masing dalam menangani tipe data tertentu.

Penerapan model klasifikasi dapat membantu mengidentifikasi tingkat kepuasan pengunjung di *Cafe Teman Nongkrong* dengan lebih sistematis. Data penilaian pengunjung yang beragam, seperti rasa makanan, harga, pelayanan, dan kenyamanan, dapat dianalisis untuk menentukan apakah seseorang termasuk kategori puas atau tidak puas. Model seperti *Decision Tree* dapat memetakan pola keputusan dengan jelas, sehingga pengelola dapat melihat faktor mana yang memiliki pengaruh paling besar terhadap kategori kepuasan. Melalui hasil klasifikasi, pengelola kafe dapat memahami karakteristik pengunjung secara lebih mendalam, memperbaiki bagian-bagian yang menimbulkan ketidakpuasan, serta merancang strategi peningkatan layanan yang lebih tepat sasaran. Model ini juga memberikan gambaran prediktif yang bisa digunakan untuk memonitor kepuasan di masa mendatang.

2.8. Alat Bantu Analisis Data

Alat bantu analisis data berperan penting dalam mempermudah proses pengolahan, pemodelan, hingga visualisasi data agar hasil yang diperoleh lebih

akurat dan mudah dipahami. Salah satu alat bantu yang banyak digunakan adalah RapidMiner, sebuah platform berbasis *graphical user interface (GUI)* yang memungkinkan pengguna melakukan berbagai proses analisis tanpa harus menuliskan kode secara manual (Rosika et al., 2024). RapidMiner menyediakan fungsi lengkap mulai dari *Preprocessing*, seleksi atribut, pemodelan *Machine learning*, evaluasi model, hingga visualisasi hasil (Abdullah et al., 2023). Keunggulan utama RapidMiner terletak pada antarmukanya yang intuitif dan alur kerja berbentuk *drag and drop*, sehingga sangat membantu dalam mempercepat proses analisis data, terutama bagi pengguna yang tidak terlalu mahir dalam pemrograman. Dengan dukungan berbagai algoritma dan operator yang fleksibel, RapidMiner mampu mengolah data berukuran besar dan menghasilkan model prediktif yang berkualitas (Ampera et al., 2024).

Penggunaan RapidMiner dapat mempermudah proses analisis tingkat kepuasan pengunjung di Cafe Teman Nongkrong karena platform ini mampu mengelola seluruh tahapan mulai dari pembersihan data, pemilihan fitur yang relevan, penerapan metode *Decision Tree*, hingga evaluasi performa model. Data penilaian pengunjung dapat dimasukkan ke dalam RapidMiner untuk kemudian diproses menggunakan operator-operator yang tersedia, sehingga seluruh prosedur analisis dapat dijalankan secara otomatis tanpa memerlukan penulisan kode secara manual. Selain itu, RapidMiner memiliki tampilan antarmuka berbasis *drag and drop* yang memudahkan dalam membangun alur kerja, mulai dari input data, transformasi, pemodelan, hingga pengujian model. Struktur pohon keputusan yang dihasilkan dapat divisualisasikan dengan jelas sehingga memudahkan dalam

mengidentifikasi faktor-faktor yang paling memengaruhi kepuasan pelanggan, serta dapat menjadi dasar bagi pengelola kafe dalam menentukan strategi peningkatan layanan, harga, maupun kualitas produk secara lebih terukur (Sudarsono & Leo, 2021).

2.9. Penelitian Terdahulu

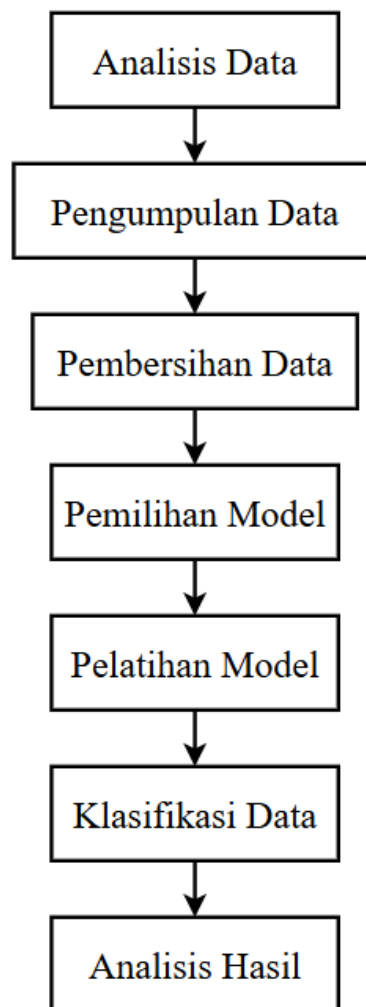
Tabel 2. 1. Penelitian Terdahulu

No	Peneliti	Judul Penelitian	Tahun	Metode dan Data	Hasil
1	(Cahyaningtyas et al., 2021)	Analisis sentimen pada rating aplikasi Shopee menggunakan metode <i>Decision Tree</i> berbasis SMOTE	2021	Metode <i>Decision Tree</i> dan data Teks Ulasan Rating Pengguna Aplikasi Shopee	Penerapan metode <i>Decision Tree</i> menghasilkan tingkat akurasi dan performa klasifikasi yang sangat tinggi baik menggunakan SMOTE maupun tanpa SMOTE. Penggunaan SMOTE terbukti meningkatkan nilai akurasi dan AUC, meskipun nilai recall dan precision tetap stabil pada kedua skenario.
2	(Azmi et al., 2023)	Analisis Pengaruh Komposisi Data Training dan Data Testing pada Penggunaan PCA dan Algoritma <i>Decision Tree</i> untuk Klasifikasi Penderita Penyakit Liver	2023	Metode <i>Decision Tree</i> dan Data yang digunakan yaitu Indian Liver Patient Dataset (ILPD) dengan 583 data	Model <i>Decision Tree</i> C4.5 yang menggunakan fitur hasil PCA mampu menghasilkan akurasi tertinggi ketika komponen PCA berjumlah 8. Distribusi data latih dan uji 90%:10% memberikan performa terbaik untuk klasifikasi penyakit liver.

3	(Aldisa & Maulana, 2022)	Analisis Sentimen Opini Masyarakat Terhadap Vaksinasi Booster COVID 19 Dengan Perbandingan Metode Naive Bayes, <i>Decision Tree</i> dan SVM	2022	Metode Naive Bayes, <i>Decision Tree</i> , SVM dan data opini Masyarakat dari twitter	Model <i>Decision Tree</i> menghasilkan performa yang lebih rendah dibandingkan SVM dan Naive Bayes berdasarkan pengukuran AUC dan presisi. Meskipun demikian, <i>Decision Tree</i> tetap mampu mengklasifikasikan sentimen dengan struktur model yang mudah dipahami.
4	(Jtik et al., 2022)	Analisis Sentimen Pembelajaran Daring menggunakan Metode Naive Bayes, KNN, dan <i>Decision Tree</i>	2022	Metode Naive Bayes, KNN, <i>Decision Tree</i> dan Data Twitter yang diperoleh melalui crawling.	Metode <i>Decision Tree</i> memperoleh performa tertinggi dengan akurasi 61.92%. Selain itu, metode ini juga menghasilkan nilai precision sebesar 73.63% dan recall sebesar 11.42%.
5	(Mardiani et al., 2023)	Komparasi Metode Knn, Naive Bayes, <i>Decision Tree</i> , Ensemble, Linear Regression Terhadap Analisis Performa Pelajar Sma	2023	Metode <i>Decision Tree</i> dan Data diperoleh dari 3 SMA yang mencakup nilai akademik dan informasi demografi siswa.	Model <i>Decision Tree</i> berhasil memetakan faktor-faktor yang paling memengaruhi performa siswa berdasarkan data nilai dan demografi. Pohon keputusan yang dihasilkan mampu mengelompokkan siswa ke dalam kategori kinerja tertentu dengan struktur aturan yang jelas dan mudah diinterpretasikan.
6	Tria syahputri	Implementasi Metode	2025	Metode Decisin	Penelitian ini menerapkan

		<i>Decision Tree</i> untuk Prediksi Tingkat Kepuasan Pengunjung di Cafe Teman Nongkrong		Tree menggunakan data	penggunaan metode decision tree untuk melakukan klasifikasi tingkat kepuasan Pengunjung di Cafe Teman Nongkrong. Penelitian ini juga ingin melihat kualitas dari metode decision tree dalam melakukan klasifikasi data.
--	--	---	--	-----------------------	---

2.10. Kerangka Kerja Penelitian



Gambar 2. 1. Kerangka Kerja Penelitian

Berikut adalah penjelasan dari masing masing poin yang ada pada kerangka kerja penelitian yaitu sebagai berikut.

1. Analisis data adalah proses menentukan variabel dan atribut yang akan digunakan untuk penelitian, sehingga data yang dipilih benar-benar relevan dengan tujuan prediksi. Tahap ini memastikan bahwa setiap variabel memiliki kontribusi penting terhadap model klasifikasi yang akan dibangun.
2. Pengumpulan data merupakan tahap mengumpulkan informasi melalui kuesioner yang disebarakan kepada pengunjung untuk mendapatkan penilaian mereka. Hasil pengisian kuesioner otomatis akan tercatat pada Google Spreadsheet sehingga memudahkan proses pengolahan data.
3. Pembersihan data adalah proses memeriksa dan menyeleksi data untuk menghilangkan kesalahan seperti nilai kosong, data tidak lengkap, atau penulisan yang typo. Tahap ini memastikan kualitas dataset tetap baik sehingga tidak mengganggu keakuratan hasil model.
4. Pemilihan model dilakukan dengan menetapkan metode *Decision Tree* sebagai teknik klasifikasi yang digunakan untuk mengolah dan memprediksi tingkat kepuasan pengunjung. Keputusan memilih model ini didasari pada kemampuan *Decision Tree* dalam memberikan struktur keputusan yang jelas dan mudah diinterpretasikan.
5. Pelatihan model adalah tahap membentuk dan menyesuaikan model *Decision Tree* dengan data yang telah dibersihkan agar algoritma dapat mempelajari pola dari variabel-variabel yang tersedia. Proses ini

menghasilkan model yang siap digunakan untuk melakukan klasifikasi pada data baru.

6. Klasifikasi data adalah proses penghitungan dan pemetaan data menggunakan model *Decision Tree* untuk menentukan kategori tingkat kepuasan pengunjung. Tahap ini menghasilkan output berupa klasifikasi apakah pengunjung berada pada kategori puas atau tidak puas berdasarkan pola yang telah dipelajari model.
7. Analisis hasil merupakan tahap meninjau kembali performa model dari hasil klasifikasi untuk melihat akurasi dan kesesuaian prediksi dengan data sebenarnya. Tahap ini juga memberikan pemahaman mengenai faktor yang paling berpengaruh terhadap tingkat kepuasan pengunjung.