

BAB II

LANDASAN TEORI

2.1. Teori-teori dan Konsep yang Mendasari Penelitian

Penelitian ini didasarkan pada teori-teori inti dalam *machine learning* yang menekankan konsep klasifikasi data dan probabilitas sebagai dasar penerapan algoritma Naïve Bayes untuk memprediksi kelarisan produk di Coffeeshop Bobskuy. Pada konteks ini, penting untuk menggali bagaimana berbagai teori mendasari desain dan penerapan algoritma klasifikasi serta relevansinya dalam analisis data penjualan produk. Pemahaman ini akan memberikan gambaran tentang bagaimana algoritma klasifikasi bekerja dan bagaimana probabilitas diterapkan untuk memprediksi hasil berdasarkan data yang ada. Penjelasan tentang teori-teori ini juga akan menunjukkan bagaimana Naïve Bayes berfungsi secara matematis dan praktis dalam memprediksi penjualan produk, serta mengapa ia merupakan pilihan yang tepat untuk tugas ini.

2.1.1. Teori Pembelajaran Statistik

Salah satu teori penting yang mendasari klasifikasi dalam pembelajaran mesin adalah Teori Pembelajaran Statistik yang dikombinasikan dengan Pemangkasan Risiko Struktural (SRM). Teori ini fokus pada bagaimana suatu model dapat meminimalkan risiko empiris, atau kesalahan yang terjadi pada data pelatihan, sambil menghindari overfitting yang dapat membatasi kemampuan model untuk menggeneralisasi ke data yang belum terlihat (Ju et al., 2021).

Pemangkasan risiko struktural formalizes keseimbangan antara kompleksitas model dan kesalahan model untuk memastikan bahwa model yang dihasilkan tidak hanya cocok dengan data yang ada, tetapi juga dapat bekerja dengan baik saat dihadapkan dengan data baru. Meskipun teori pembelajaran statistik juga mendasari algoritma lain seperti Support Vector Machine, penelitian ini memilih Naïve Bayes karena kemampuannya menghasilkan prediksi cepat dan mudah diinterpretasikan, yang sesuai untuk kebutuhan analisis penjualan di Coffeeshop Bobskuy yang lebih sesuai untuk aplikasi di lingkungan yang membutuhkan keputusan cepat, seperti dalam prediksi penjualan produk.

2.1.2. Teori Keputusan Bayesian

Teori berikutnya yang mendasari *Naïve Bayes* adalah Teori Keputusan Bayesian dan Klasifikasi Probabilistik, yang membentuk dasar bagi klasifikasi probabilistik. Naïve Bayes menggunakan Teorema Bayes, yang memungkinkan model untuk menghitung probabilitas posterior untuk setiap kelas berdasarkan fitur yang diamati, dengan menggunakan probabilitas sebelumnya dan kemungkinan fitur yang relevan (Wibowo et al., 2023). Dalam hal ini, probabilitas kelas dihitung untuk setiap kemungkinan hasil, dan kelas dengan probabilitas tertinggi dipilih sebagai keputusan klasifikasi. Keunggulan utama dari Naïve Bayes adalah asumsi independensi kondisional yang menyederhanakan perhitungan probabilitas dan memungkinkan pelatihan model yang cepat dan estimasi yang efisien meskipun dengan data yang memiliki dimensi tinggi. Naïve Bayes tetap efektif, bahkan dalam situasi di mana fitur-fitur yang digunakan tidak sepenuhnya independen satu sama lain (Punia et al., 2021). Hal ini menjadikan *Naïve Bayes* sebagai model yang sangat

berguna dalam prediksi penjualan produk di industri retail, di mana sejumlah besar fitur perlu dianalisis untuk menentukan permintaan produk atau preferensi pelanggan.

2.1.3. Teori Informasi

Teori Informasi berperan penting dalam proses klasifikasi untuk mengukur relevansi antar fitur melalui konsep mutual information yang membantu menentukan atribut penjualan produk paling berpengaruh terhadap tingkat kelarisan pada Coffeeshop Bobskuy. Dalam pembelajaran mesin, informasi mutual digunakan untuk mengukur sejauh mana dua variabel terkait satu sama lain, yang berguna untuk memilih fitur yang memiliki relevansi tertinggi terhadap klasifikasi. Penggunaan teori informasi ini membantu dalam membuat keputusan mengenai fitur mana yang perlu diprioritaskan untuk menghasilkan model yang lebih akurat dan efisien. Dalam konteks prediksi penjualan, teori informasi ini dapat membantu memilih atribut produk atau data pelanggan yang paling berpengaruh dalam menentukan permintaan pasar.

2.1.4. Teori Kategori

Teori Kategori memberikan wawasan tentang hubungan dan transformasi dalam pembelajaran mesin. Dalam hal ini, teori kategori membantu mengatur proses pembelajaran dalam model klasifikasi dan memberikan kerangka kerja yang lebih luas untuk memahami interaksi antara berbagai elemen dalam model pembelajaran mesin (Chen et al., 2021). Dengan teori ini, algoritma pembelajaran mesin secara lebih sistematis, menyatukan berbagai pendekatan klasifikasi untuk mencapai hasil yang optimal.

Meskipun Naïve Bayes menawarkan kemudahan dan kecepatan dalam klasifikasi, tidak semua data cocok dengan asumsi independensi yang mendasari model ini. Oleh karena itu, pemilihan fitur yang tepat sangat penting untuk memastikan bahwa Naïve Bayes memberikan hasil yang optimal. Salah satu cara untuk mengatasi keterbatasan ini adalah dengan menggunakan teknik seleksi fitur yang efektif, yang memungkinkan model untuk memfokuskan hanya pada fitur-fitur yang paling relevan (Wibowo et al., 2023). Dalam prediksi penjualan produk, pemilihan fitur yang cermat memungkinkan model untuk lebih efisien dalam menganalisis data, mengurangi kemungkinan kesalahan dan meningkatkan akurasi prediksi.

Dalam prakteknya, Naïve Bayes telah terbukti efektif dalam berbagai aplikasi, terutama dalam analisis retail. Misalnya, di bidang pemasaran, Naïve Bayes digunakan untuk mengklasifikasikan produk berdasarkan permintaan yang diharapkan, melakukan segmentasi pelanggan, dan merencanakan strategi promosi. Model ini sangat cocok untuk digunakan dalam situasi di mana waktu keputusan sangat penting, seperti merespon perubahan cepat dalam permintaan pasar atau dalam merencanakan promosi produk secara real-time (Punia et al., 2021); (Wahid et al., 2023). Kecepatan dan interpretabilitas Naïve Bayes membuatnya sangat berguna dalam konteks bisnis yang membutuhkan kecepatan tanpa mengorbankan akurasi.

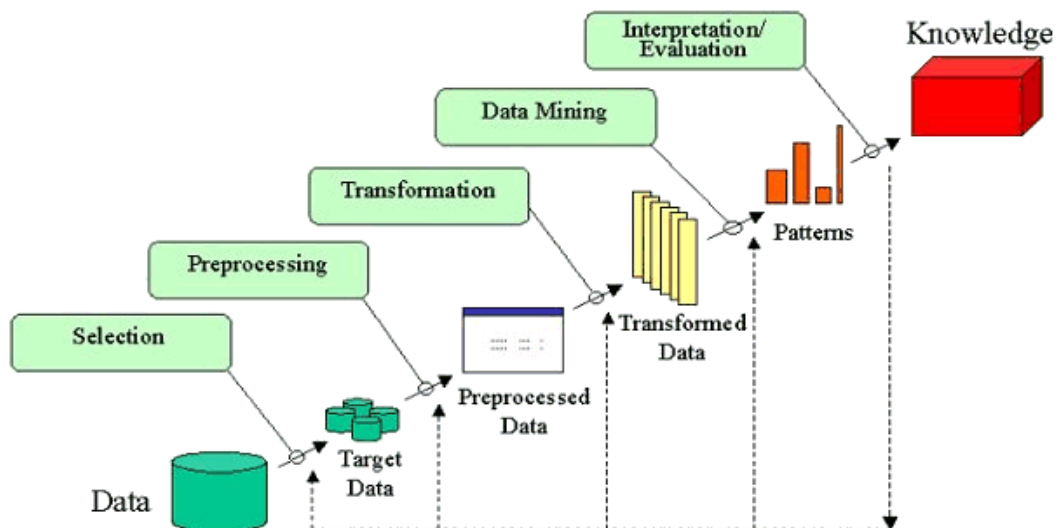
Meskipun Naïve Bayes memiliki banyak kelebihan, ia juga memiliki keterbatasan praktis, terutama terkait dengan asumsi independensi yang mungkin tidak selalu sesuai dengan data yang ada. Oleh karena itu, perlu dilakukan pemilihan

fitur yang cermat dan validasi model terhadap metode lain untuk memastikan bahwa hasil yang diperoleh tetap akurat. Dalam beberapa kasus, model yang lebih kompleks seperti SVM atau jaringan saraf tiruan dapat digunakan untuk meningkatkan akurasi klasifikasi, meskipun dengan biaya komputasi yang lebih tinggi (Wang et al., 2023).

2.2. Knowledge Discovery in Databases (KDD)

Knowledge Discovery in Databases (KDD) merupakan kerangka kerja konseptual yang menjelaskan proses sistematis dalam mengekstraksi pengetahuan yang berguna dari kumpulan data berukuran besar. KDD tidak merujuk pada satu teknik tunggal, melainkan rangkaian tahapan terstruktur yang mencakup pemilihan data, pembersihan data, transformasi data, penerapan metode analitik, hingga interpretasi hasil.

2.2.1. Tahapan Knowledge Discovery in Databases (KDD)



Gambar 2. 1 Proses KDD

a. Selection (Seleksi Data)

Tahap seleksi data merupakan langkah awal dalam proses KDD yang

berfokus pada pemilihan data yang relevan dengan tujuan analisis. Tidak semua data yang tersedia dalam basis data digunakan, melainkan hanya atribut atau variabel yang memiliki keterkaitan langsung dengan permasalahan penelitian. Proses ini bertujuan untuk meningkatkan efisiensi analisis dan mengurangi kompleksitas pemodelan. Dalam konteks penelitian penjualan, seleksi data melibatkan identifikasi variabel penting seperti nama produk, harga, jumlah terjual, dan total pendapatan yang akan digunakan dalam proses klasifikasi.

b. Preprocessing (Praproses Data)

Tahap praproses data bertujuan untuk membersihkan data dari berbagai ketidaksempurnaan yang dapat memengaruhi kualitas model. Kegiatan pada tahap ini mencakup penanganan data hilang (missing values), penghapusan data duplikat, koreksi inkonsistensi, serta eliminasi noise. Data yang tidak bersih dapat menyebabkan bias dan menurunkan akurasi model data mining.

c. Transformation (Transformasi Data)

Transformasi data merupakan proses pengubahan atau penyesuaian format data agar sesuai dengan kebutuhan algoritma yang digunakan. Tahapan ini dapat mencakup normalisasi, diskretisasi, agregasi, maupun konstruksi fitur baru. Tujuan utama transformasi adalah meningkatkan representasi data sehingga pola yang tersembunyi dapat lebih mudah diidentifikasi oleh model.

d. Data Mining

Data mining adalah tahap inti dalam proses KDD, di mana teknik

komputasional diterapkan untuk menemukan pola, hubungan, atau model prediktif dari data yang telah diproses. Pada tahap ini, algoritma machine learning digunakan untuk melakukan tugas analitik seperti klasifikasi, klusterisasi, asosiasi, atau prediksi. Pemilihan metode data mining sangat bergantung pada tujuan penelitian. Dalam kasus klasifikasi penjualan, algoritma seperti Naïve Bayes digunakan untuk membangun model probabilistik yang mampu mengelompokkan data ke dalam kategori tertentu berdasarkan pola historis.

e. **Interpretation / Evaluation (Interpretasi dan Evaluasi)**

Tahap interpretasi dan evaluasi bertujuan untuk menilai kualitas serta validitas pola atau model yang dihasilkan pada tahap data mining. Tidak semua pola yang ditemukan memiliki nilai pengetahuan yang berguna, sehingga diperlukan evaluasi untuk menentukan relevansi dan akurasi hasil. Evaluasi dapat dilakukan menggunakan metrik kinerja seperti akurasi, presisi, recall, dan F1-score.

f. **Knowledge (Pengetahuan)**

Tahap akhir KDD adalah pembentukan pengetahuan, yaitu hasil akhir berupa informasi yang bermakna dan dapat digunakan secara praktis. Pengetahuan yang diperoleh dapat berupa aturan, model prediktif, atau wawasan strategis yang mendukung proses pengambilan keputusan. Dalam konteks bisnis, pengetahuan ini dapat dimanfaatkan untuk merumuskan strategi pemasaran, pengelolaan stok

2.3. Langkah-Langkah Data Mining

Data mining merupakan proses analisis data yang bertujuan untuk mengidentifikasi pola, tren, serta informasi penting dari kumpulan data yang berukuran besar dan kompleks.

2.3.1. Pengumpulan Data

Pengumpulan data merupakan langkah awal yang sangat penting dalam proses data mining. Dalam penelitian ini, data yang digunakan diperoleh dari transaksi penjualan produk. Tahap pengumpulan data ini sangat krusial, karena kualitas data yang diperoleh akan mempengaruhi hasil klasifikasi dan prediksi yang dilakukan menggunakan algoritma Naïve Bayes. Oleh karena itu, memastikan bahwa data yang terkumpul adalah relevan, akurat, dan lengkap akan sangat menentukan efektivitas model yang dikembangkan untuk mendukung pengambilan keputusan berbasis data.

2.3.2. Pembersihan Data (Data Cleaning)

Langkah selanjutnya adalah pembersihan data. Tahap ini bertujuan untuk memastikan bahwa data yang akan digunakan dalam analisis berada dalam kondisi yang bersih, bebas dari kesalahan, dan konsisten. Dalam praktiknya, data mentah sering kali mengandung berbagai masalah, seperti nilai yang hilang (missing values), duplikasi, kesalahan pengetikan, atau data yang tidak relevan. Jika masalah ini tidak ditangani dengan baik, algoritma Naïve Bayes yang digunakan dalam

penelitian ini bisa salah dalam membaca pola, yang pada akhirnya menghasilkan klasifikasi yang tidak akurat. Oleh karena itu, tahap pembersihan data ini sangat penting dan menjadi dasar yang kokoh sebelum melanjutkan ke tahap analisis lebih lanjut.

2.3.3. Transformasi Data

Transformasi data adalah tahap di mana data yang telah dibersihkan diubah ke dalam format yang sesuai dengan kebutuhan algoritma yang digunakan. Karena tidak semua algoritma dapat langsung memproses data dalam bentuk aslinya, maka data perlu melalui proses normalisasi, kategorisasi, atau pengkodean ulang. Dalam penelitian ini, data penjualan yang telah dibersihkan kemudian ditransformasikan ke dalam bentuk tabel probabilitas yang sesuai dengan metode Naïve Bayes.

2.3.4. Penerapan Algoritma Data Mining

Tahap inti dalam proses data mining adalah penerapan algoritma. Pada tahap ini, data yang telah diproses dan disiapkan sebelumnya akan diproses menggunakan metode yang sesuai dengan tujuan analisis yang ingin dicapai. Algoritma yang digunakan dalam penelitian ini adalah algoritma klasifikasi, yaitu Naïve Bayes, yang diterapkan untuk mengklasifikasikan produk di Coffeeshop Bobskuy sebagai produk laris atau tidak laris. Pada tahap ini, data training digunakan untuk membangun model klasifikasi, sementara data testing digunakan untuk mengevaluasi kinerja model yang telah dibangun, guna memastikan bahwa model dapat memprediksi dengan akurat berdasarkan data historis penjualan.

2.3.5. Evaluasi Dan Interpretasi

Setelah algoritma diterapkan, langkah berikutnya adalah evaluasi hasil. Evaluasi ini bertujuan untuk menilai seberapa baik model yang dibangun dalam mengklasifikasikan atau memprediksi data baru. Jika model menunjukkan akurasi yang tinggi dan memiliki keseimbangan yang baik antara presisi dan recall, maka model tersebut dianggap valid dan siap digunakan. Namun, jika kinerjanya rendah, maka perlu dilakukan perbaikan, baik pada tahap pembersihan data, transformasi data, atau pemilihan algoritma yang lebih tepat. Evaluasi ini juga bertujuan untuk memastikan bahwa hasil klasifikasi tidak hanya akurat dari segi perhitungan matematis, tetapi juga relevan secara praktis sesuai dengan konteks permasalahan yang dihadapi. Interpretasi hasil sangat penting agar temuan dari proses data mining dapat dipahami dengan jelas dan dapat digunakan sebagai dasar dalam pengambilan keputusan yang lebih tepat di Coffeeshop Bobskuy.

2.4. Algoritma yang Digunakan

2.4.1. Pengertian Naive Bayes

Algoritma *Naive Bayes* adalah metode klasifikasi probabilistik yang menggunakan Teorema Bayes dengan asumsi independensi antar fitur. Dalam penelitian ini, algoritma tersebut digunakan untuk menentukan apakah suatu produk pada Coffeeshop Bobskuy termasuk kategori laris atau tidak laris berdasarkan atribut penjualan. Dalam konteks penelitian ini, algoritma *Naive Bayes* digunakan untuk memprediksi apakah suatu produk di Coffeeshop Bobskuy termasuk dalam

kategori laris atau tidak laris berdasarkan data penjualan. Pendekatan ini dipilih karena kemampuannya yang kuat dalam menangani klasifikasi dengan data yang bersifat campuran serta efisiensinya dalam proses komputasi.

2.4.2. Prinsip Dasar Algoritma Naïve Bayes

Prinsip dasar algoritma Naïve Bayes didasarkan pada Teorema Bayes, yaitu suatu pendekatan probabilistik yang digunakan untuk menghitung peluang suatu data termasuk ke dalam kelas tertentu berdasarkan fitur-fitur yang dimilikinya (Bishnoi et al., 2022).

Rumus dasar pada perhitungan Naïve Bayes:

$$P(C | X) = \frac{P(X | C) \cdot P(C)}{P(X)}$$

$P(C|X)$ adalah probabilitas data X terhadap kelas C , $P(X|C)$ adalah probabilitas fitur X pada kelas C , $P(C)$ adalah probabilitas awal suatu kelas, dan $P(X)$ adalah probabilitas total dari data X . Hasil klasifikasi ditentukan berdasarkan nilai probabilitas terbesar dari masing-masing kelas.

Dalam rumus Naïve Bayes, C merupakan kelas atau kategori yang menjadi tujuan klasifikasi. Pada penelitian ini, kelas C terdiri dari dua kategori, yaitu Laris dan Tidak Laris. Sementara itu, X merupakan data yang akan diklasifikasikan, yaitu sekumpulan atribut yang dimiliki oleh suatu produk. Adapun $P(X|C)$ merupakan probabilitas munculnya fitur atau atribut X pada kelas C . Artinya, nilai ini menunjukkan seberapa besar kemungkinan suatu data dengan karakteristik tertentu muncul dalam kelas tertentu. Sebagai contoh, $P(X|Laris)$ menunjukkan peluang suatu produk dengan atribut tertentu muncul pada kelas Laris, sedangkan $P(X|Tidak$

Laris) menunjukkan peluang produk dengan atribut tersebut muncul pada kelas Tidak Laris.

Pemilihan algoritma ini juga didasarkan pada karakteristiknya yang sederhana namun efektif. Seperti dijelaskan oleh (Abobakir & Abdulazeez, 2024), *Naïve Bayes* merupakan salah satu algoritma klasifikasi tertua dan paling banyak digunakan dalam pembelajaran mesin karena efisiensi komputasinya yang tinggi serta kemampuannya memberikan hasil yang akurat pada berbagai jenis data. Pendekatan ini juga memberikan keluaran berupa probabilitas, sehingga pengguna tidak hanya mengetahui prediksi kelas tetapi juga tingkat keyakinan terhadap prediksi tersebut, yang sangat relevan dalam pengambilan keputusan bisnis.

Naïve Bayes menunjukkan kemampuan adaptasi yang tinggi terhadap data campuran, baik yang bersifat numerik maupun kategorikal. (Bishnoi et al., 2022) menunjukkan bahwa algoritma ini dapat diterapkan pada dataset yang mengandung variabel kontinyu dan kategorikal secara bersamaan, dengan hasil yang kompetitif dibandingkan model lain seperti Decision Tree atau K-Nearest Neighbors. Hal ini sangat relevan dalam penelitian di Coffeeshop Bobskuy.

2.4.3. Keunggulan Naïve Bayes

Keunggulan *Naïve Bayes* pada kecepatan pelatihan dan inferensi membuatnya ideal untuk sistem prediksi penjualan yang memerlukan pembaruan data secara berkala. Dalam studi yang dilakukan oleh (Saputri et al., 2024), algoritma ini terbukti efisien untuk analisis sentimen dengan jumlah data besar karena kesederhanaan struktur perhitungannya. Kesamaan sifat antara data teks dan

data penjualan yang tinggi dimensinya menunjukkan bahwa pendekatan ini juga sesuai untuk kasus prediksi penjualan.

Selain cepat dan efisien, *Naïve Bayes* memiliki interpretabilitas yang tinggi karena hasil perhitungannya menunjukkan pengaruh setiap variabel, seperti harga dan promosi, terhadap peluang suatu produk menjadi laris di Coffeeshop Bobskuy. Probabilitas posterior yang dihasilkan memberikan gambaran yang jelas mengenai pengaruh masing-masing fitur terhadap hasil klasifikasi. Dalam konteks bisnis, hal ini penting untuk memahami faktor-faktor yang paling berpengaruh terhadap tingkat kelarisan produk. (Boshoma et al., 2025) menekankan bahwa kemampuan algoritma ini dalam memberikan interpretasi berbasis probabilistik membuatnya cocok untuk pengambilan keputusan yang membutuhkan pemahaman risiko dan peluang.

Namun demikian, asumsi independensi antar fitur tetap menjadi keterbatasan utama algoritma ini. (Aman & Chhillar, 2023) untuk mengatasi masalah ini dengan menggabungkan *Naïve Bayes* dalam pendekatan ensemble, khususnya melalui metode stacking, di mana hasil prediksi *Naïve Bayes* digunakan bersama algoritma lain seperti *Logistic Regression* dan *Support Vector Machine*. Pendekatan ini terbukti meningkatkan performa prediksi pada kasus kompleks seperti diagnosis diabetes, yang secara karakteristik serupa dengan prediksi penjualan produk karena melibatkan data campuran.

Penggunaan *Naïve Bayes* juga memiliki keterkaitan kuat dengan penelitian (Lubis et al., 2024) yang membandingkan performa algoritma *Naïve Bayes*, *Decision Tree*, dan *K-Nearest Neighbors* pada klasifikasi pola makan penderita

diabetes. Hasil penelitian tersebut menunjukkan bahwa meskipun asumsi independensi dapat membatasi akurasi dalam kasus tertentu, *Naïve Bayes* tetap menunjukkan hasil yang kompetitif karena kemampuannya menangani data dalam jumlah besar dengan kecepatan tinggi. Implikasi ini relevan untuk Coffeeshop Bobskuy yang mungkin menghadapi dataset penjualan dalam volume besar dan perlu melakukan analisis secara cepat untuk mendukung strategi pemasaran.

Selain sebagai model mandiri, *Naïve Bayes* juga sering digunakan sebagai komponen dasar dalam model hibrida. (Aman & Chhillar, 2023) dan menunjukkan bahwa kombinasi *Naïve Bayes* dengan model lain seperti *Random Forest* dan *Artificial Neural Network* dapat menghasilkan peningkatan signifikan dalam akurasi prediksi. Pendekatan ini memperluas fleksibilitas *Naïve Bayes* untuk diterapkan pada konteks data yang lebih kompleks tanpa kehilangan keunggulan dasarnya, yaitu interpretabilitas dan efisiensi.

(Wijaya et al., 2022) juga menunjukkan efektivitas *Naïve Bayes* dalam konteks pendidikan untuk memprediksi tingkat keberhasilan belajar mahasiswa. Proses klasifikasi yang digunakan—menghitung prior, likelihood, dan posterior untuk setiap kelas—menunjukkan paralelisme dengan konteks penelitian ini, di mana prediksi kelulusan produk memerlukan proses penghitungan probabilitas yang serupa. Temuan tersebut memperkuat argumen bahwa *Naïve Bayes* dapat digunakan dalam berbagai domain yang memerlukan pengambilan keputusan berbasis probabilitas.

Kelebihan utama dari algoritma ini dalam konteks penjualan terletak pada kemampuannya menghasilkan prediksi probabilistik yang dapat diinterpretasikan.

Hal ini sangat berguna dalam perencanaan stok dan strategi promosi, karena setiap produk dapat dinilai berdasarkan tingkat probabilitas kelarisan. Misalnya, jika suatu produk memiliki probabilitas tinggi untuk menjadi laris, Coffeeshop Bobskuy dapat menyesuaikan strategi persediaan untuk memenuhi permintaan. Sebaliknya, produk dengan probabilitas rendah dapat dipertimbangkan untuk promosi atau penyesuaian harga.

Meskipun demikian, kelemahan algoritma ini tidak dapat diabaikan. Seperti yang dijelaskan oleh (Bishnoi et al., 2022), korelasi antar fitur dapat menyebabkan bias dalam estimasi probabilitas jika asumsi independensi tidak terpenuhi. Dalam konteks penjualan, fitur seperti cuaca, promosi, dan harga mungkin berhubungan erat, sehingga pendekatan alternatif seperti *Random Forest* atau *model ensemble* dapat digunakan untuk meningkatkan hasil prediksi. (Aman & Chhillar, 2023) menunjukkan bahwa integrasi *Naïve Bayes* ke dalam kerangka *ensemble* meningkatkan ketahanan model terhadap masalah korelasi antar fitur.

Selain itu, (Boshoma et al., 2025) menekankan bahwa performa *Naïve Bayes* sangat bergantung pada bentuk distribusi data. Untuk fitur kontinu, pendekatan *Gaussian Naïve Bayes* biasanya digunakan dengan asumsi bahwa data mengikuti distribusi normal. Namun, jika distribusi aktual data berbeda secara signifikan, maka hasil prediksi dapat terpengaruh. Dalam penelitian ini, strategi yang dapat diterapkan untuk mengatasi masalah tersebut adalah melakukan diskretisasi atau normalisasi fitur kontinu agar lebih sesuai dengan asumsi model.

Dalam kerangka penerapan di Coffeeshop Bobskuy, *Naïve Bayes* dapat diintegrasikan sebagai bagian dari sistem prediksi berbasis data historis. Dengan

memanfaatkan data transaksi, algoritma ini akan menghitung probabilitas setiap kategori kelarisan berdasarkan atribut produk dan konteks penjualannya. Proses prediksi ini meliputi empat tahap utama sebagaimana diuraikan oleh (Saputri et al., 2024): (1) perhitungan prior berdasarkan proporsi kelas dalam data pelatihan, (2) estimasi likelihood untuk setiap fitur terhadap kelas tertentu, (3) perhitungan posterior untuk setiap kelas berdasarkan data baru, dan (4) pengambilan keputusan dengan memilih kelas yang memiliki probabilitas tertinggi.

Keterbatasan asumsi independensi dapat diatasi dengan menerapkan pendekatan ensemble atau kombinasi dengan model lain. (Aman & Chhillar, 2023) membuktikan bahwa *Naïve Bayes* sebagai base learner dalam model stacking mampu memberikan kontribusi signifikan terhadap peningkatan performa sistem prediksi secara keseluruhan. Hal ini membuka peluang untuk menerapkan strategi serupa di Coffeeshop Bobskuy, di mana *Naïve Bayes* dapat dikombinasikan dengan model seperti Decision Tree atau Logistic Regression untuk mendapatkan hasil yang lebih robust.

Secara keseluruhan, *Naïve Bayes* tetap menjadi algoritma yang sangat relevan dan efektif untuk penelitian ini. Dengan dasar teoritis yang kuat dalam Teorema Bayes, efisiensi komputasi, serta kemampuannya menghasilkan prediksi probabilistik yang mudah diinterpretasikan, algoritma ini memberikan landasan metodologis yang solid bagi sistem prediksi penjualan Coffeeshop Bobskuy. Kelebihannya dalam kecepatan, kesederhanaan, dan skalabilitas menjadikannya solusi ideal untuk data penjualan berskala besar. kombinasi dengan pendekatan

ensemble dapat menjadi solusi praktis yang mempertahankan keunggulan *Naïve Bayes* sambil meningkatkan akurasinya dalam konteks nyata.

2.5. Langkah-Langkah Machine Learning

Proses pembelajaran mesin merupakan serangkaian tahapan yang saling berkaitan untuk membangun model yang mampu melakukan prediksi atau klasifikasi berdasarkan data. Dalam konteks penelitian ini, tahapan-tahapan tersebut sangat penting untuk memahami bagaimana data diproses sebelum digunakan dalam algoritma *Naïve Bayes* untuk memprediksi kelarisan produk di Coffeeshop Bobskuy. Setiap tahap dalam proses *machine learning* berperan penting untuk menghasilkan model prediksi yang akurat dan andal dalam menentukan kelarisan produk pada Coffeeshop Bobskuy.

2.5.1. Pengumpulan Data

Tahapan pertama dalam proses *machine learning* adalah pengumpulan data. Data merupakan bahan utama yang menentukan kualitas hasil model prediktif. Sebagaimana dinyatakan oleh (Tanjung et al., 2023), setiap implementasi algoritma *Naïve Bayes* dimulai dengan tahapan pengumpulan data yang sistematis untuk memastikan bahwa informasi yang diperoleh relevan dengan tujuan analisis. Pengumpulan data dapat dilakukan melalui transaksi penjualan, ulasan pelanggan, atau sistem manajemen inventori. Dalam konteks ini, framing masalah menjadi langkah awal penting untuk menentukan parameter apa yang akan digunakan dalam model.

2.5.2. Praproses Data (Data Preprocessing)

Setelah data terkumpul, langkah berikutnya adalah praproses data (data preprocessing). Proses ini merupakan tahap krusial yang memastikan bahwa data yang digunakan bersih, terstruktur, dan sesuai untuk pelatihan model. (Meidina & Abidin, 2023) menekankan bahwa tahap pembersihan data melibatkan penghapusan noise, penanganan nilai hilang, serta koreksi terhadap data yang tidak konsisten. Dalam penelitian, proses pembersihan data pada algoritma *Naïve Bayes* berperan penting dalam meningkatkan akurasi prediksi penyakit jantung. Tahapan ini relevan dengan penelitian di Coffeeshop Bobskuy, di mana data transaksi dapat mengandung kesalahan input atau data duplikat yang perlu diatasi sebelum model dapat dilatih.

2.5.3. Normalisasi dan Diskretisasi Data

Normalisasi dan diskretisasi dilakukan untuk menyesuaikan format data dengan kebutuhan algoritma. (Meidina & Abidin, 2023) juga menyoroti pentingnya normalisasi, terutama ketika variabel memiliki rentang nilai yang berbeda, agar model dapat memperlakukan setiap fitur secara proporsional. Untuk algoritma *Naïve Bayes*, proses diskretisasi—mengubah data kontinu menjadi kategori—membantu menyesuaikan distribusi data dengan asumsi probabilistik model. Proses ini menjadikan fitur numerik seperti harga atau jumlah penjualan lebih mudah dikelola dalam perhitungan probabilitas.

2.5.4. Pengkodean Variabel Kategorikal

Pengkodean Variabel Kategorikal (Encoding) Setelah data dinormalisasi, tahap berikutnya adalah pengkodean variabel kategorikal (encoding). Menurut (Debora et al., 2024), transformasi kategori menjadi representasi numerik merupakan bagian penting dari pipeline *Naïve Bayes* karena algoritma ini memerlukan data dalam bentuk angka untuk menghitung probabilitas setiap fitur terhadap kelas.

2.5.5. Seleksi Fitur (Feature Selection)

Langkah selanjutnya adalah seleksi fitur (feature selection). Pemilihan fitur yang relevan membantu mengurangi dimensi data dan meningkatkan kinerja model. (Meidina & Abidin, 2023) menunjukkan bahwa penggunaan metode *Particle Swarm Optimization* (PSO) dan *Gain Ratio* dalam pemilihan fitur dapat meningkatkan akurasi algoritma *Naïve Bayes* dalam diagnosis penyakit jantung. Hal ini sejalan dengan temuan (Wibowo et al., 2023), yang menjelaskan bahwa pengurangan dimensi melalui metode hybrid top-k feature selection meningkatkan performa klasifikasi data berdimensi tinggi. Pemilihan fitur yang tepat membantu algoritma *Naïve Bayes* menghasilkan prediksi yang lebih akurat dan efisien

2.5.6. Augmentasi Data

Pada kasus tertentu, jika data terbatas atau tidak seimbang, proses augmentasi data (data augmentation) dapat diterapkan. (Seiler et al., 2024) menyatakan bahwa pembuatan data sintetis dapat meningkatkan keandalan model dengan memperluas variasi data pelatihan tanpa menambah biaya pengumpulan data baru. Dalam konteks Coffeeshop Bobskuy, teknik ini dapat digunakan untuk

menyeimbangkan jumlah data antara produk yang laris dan tidak laris sehingga model tidak bias terhadap kategori tertentu.

2.5.7. Pembagian Data (*Training dan Testing Set*)

Tahapan berikutnya adalah pembagian data menjadi data pelatihan dan data pengujian. Pembagian ini memungkinkan model untuk dilatih pada sebagian data dan dievaluasi pada data lain yang belum pernah dilihat sebelumnya. (Tanjung et al., 2023) menegaskan bahwa pembagian data merupakan langkah esensial dalam pipeline *Naïve Bayes* karena validasi model hanya dapat dilakukan jika data uji benar-benar independen dari data pelatihan. Demikian pula, (Vaerenberg et al., 2024) menambahkan bahwa proses validasi harus dirancang sedemikian rupa agar dapat mengevaluasi kinerja model terhadap distribusi data yang tidak terlihat (*unseen data*). Hal ini memastikan bahwa model memiliki kemampuan generalisasi yang baik ketika diterapkan pada data baru.

2.5.8. Pelatihan Model (*Model Training*)

Dalam pelatihan model, algoritma *Naïve Bayes* bekerja dengan menghitung probabilitas prior dan likelihood dari setiap fitur terhadap kelas yang ditentukan. Setelah pelatihan selesai, model diuji menggunakan data uji untuk mengukur kinerja dan tingkat akurasi. (Debora et al., 2024) menggambarkan alur ini secara jelas dalam studi mereka, di mana proses pelatihan *Naïve Bayes* diikuti oleh validasi untuk mengukur efektivitas model dalam menganalisis sentimen publik. Evaluasi hasil pengujian dilakukan dengan menghitung metrik seperti akurasi, presisi, recall, dan F1-score untuk menilai keseimbangan performa antara prediksi benar dan salah. (Pradana & Sugiharti, 2023)

2.5.9. Validasi dan Evaluasi Model

Pada tahap pengujian dan validasi model, penting untuk memastikan bahwa model tidak hanya cocok untuk data pelatihan, tetapi juga memiliki kemampuan generalisasi yang baik terhadap data baru. Menurut (Vaerenberg et al., 2024), validasi silang (cross-validation) berperan sebagai metode yang kuat untuk memperkirakan kinerja model pada data yang tidak terlihat. Dengan membagi dataset menjadi beberapa lipatan (fold), setiap lipatan bergantian digunakan sebagai data uji, sehingga diperoleh estimasi performa model yang lebih stabil dan dapat diandalkan. Pendekatan ini membantu meminimalkan overfitting dan memberikan gambaran yang lebih realistis tentang kemampuan model ketika diterapkan di dunia nyata

2.5.10. Tantangan Pelatihan Model

Tantangan yang sering muncul dalam proses pelatihan model juga perlu diperhatikan. (Mahmud et al., 2024) menyoroti bahwa ketersediaan data yang berlabel sering kali menjadi kendala pada pengembangan sistem pembelajaran mesin. Pengumpulan dan pelabelan data memerlukan waktu dan biaya yang tinggi, sehingga dalam beberapa kasus diperlukan teknik transfer pengetahuan atau pembelajaran semi-supervised untuk memanfaatkan data yang belum berlabel. Di sisi lain, (Qian, 2024) menunjukkan bahwa pemilihan dan penyetelan hyperparameter merupakan aspek kompleks yang dapat memengaruhi performa model secara signifikan. Dalam konteks *Naïve Bayes*, meskipun model ini relatif sederhana dan tidak banyak memerlukan hyperparameter, tahap penyetelan seperti

pemilihan varian Gaussian atau Multinomial tetap berpengaruh terhadap hasil akhir.

(Lin et al., 2022) mengemukakan bahwa sistem pembelajaran mesin rentan terhadap ancaman seperti pencurian model dan manipulasi data, sehingga tahap pra-proses dan pelatihan harus dirancang dengan mempertimbangkan aspek keamanan. (Xie et al., 2021) menambahkan bahwa pendekatan terdistribusi dan efisiensi komunikasi dalam sistem pembelajaran dapat membantu mengurangi risiko kebocoran data serta meningkatkan keandalan infrastruktur pelatihan.

2.5.11. Mengevaluasi Hasil Dan Menerapkan Model Naïve Bayes

Langkah terakhir adalah mengevaluasi hasil dan menerapkan model Naïve Bayes untuk merekomendasikan strategi bisnis di Coffeeshop Bobskuy. Hasil klasifikasi digunakan untuk memperkirakan produk mana yang berpotensi laris sehingga dapat dijadikan dasar penentuan stok dan promosi. Dalam konteks Coffeeshop Bobskuy, model yang telah dilatih dengan algoritma *Naïve Bayes* dapat digunakan untuk memprediksi produk mana yang berpotensi laris berdasarkan data penjualan sebelumnya. Evaluasi berkelanjutan juga diperlukan untuk memastikan bahwa model tetap relevan terhadap perubahan tren penjualan dan preferensi pelanggan.

2.6. Teknik Evaluasi Algoritma di Machine Learning

Evaluasi algoritma Naïve Bayes bertujuan menilai sejauh mana model mampu mengenali pola penjualan produk di Coffeeshop Bobskuy dan menggeneralisasi hasilnya pada data penjualan baru. Dalam konteks penelitian prediksi kelulusan produk di Coffeeshop Bobskuy dengan algoritma *Naïve Bayes*,

fokus evaluasi diarahkan pada empat metrik utama—akurasi, presisi, recall, dan F1-score—serta dukungan metrik lain seperti AUC-ROC, laporan per-kelas, dan matriks kebingungan. Pendekatan multi-metrik ini penting karena tidak ada satu metrik pun yang dapat menangkap seluruh dimensi kinerja model, terlebih pada data yang mungkin tidak seimbang antar kelas dan berisiko memunculkan biaya salah klasifikasi yang asimetris antara kesalahan positif dan negatif. Literatur lintas domain secara konsisten menegaskan bahwa akurasi kerap dilaporkan sebagai indikator awal, tetapi pemaknaannya mesti diperkaya dengan presisi, recall, dan F1 untuk mengungkap karakteristik prediksi pada tiap kelas dan implikasi risikonya.

2.6.1. Confusion Matrix

Confusion Matrix adalah tabel yang digunakan untuk mengukur kinerja model klasifikasi dengan membandingkan hasil prediksi model dengan kondisi sebenarnya. Melalui Confusion Matrix, dapat diketahui jumlah prediksi yang benar maupun yang salah dari masing-masing kelas. Dalam penelitian ini, kelas yang digunakan adalah Laris dan Tidak Laris.

1. True Positive (TP), yaitu jumlah data yang diprediksi Laris dan kenyataannya memang Laris.
2. True Negative (TN), yaitu jumlah data yang diprediksi Tidak Laris dan kenyataannya memang Tidak Laris.
3. False Positive (FP), yaitu jumlah data yang diprediksi Laris, tetapi kenyataannya Tidak Laris.

4. False Negative (FN), yaitu jumlah data yang diprediksi Tidak Laris, tetapi kenyataannya Laris.

Dengan menggunakan Confusion Matrix, peneliti dapat mengetahui pola kesalahan klasifikasi yang dihasilkan oleh model Naïve Bayes. Selain itu, Confusion Matrix juga menjadi dasar dalam menghitung nilai akurasi, presisi, recall, dan F1-score.

2.6.2. Akurasi (Accuracy)

Akurasi adalah ukuran yang digunakan untuk mengetahui tingkat ketepatan model dalam mengklasifikasikan data secara keseluruhan. Nilai akurasi menunjukkan perbandingan antara jumlah prediksi yang benar dengan seluruh jumlah data yang diuji. Semakin tinggi nilai akurasi, maka semakin baik kemampuan model dalam melakukan klasifikasi. (Mahmud et al., 2024).

Rumus akurasi adalah sebagai berikut:

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

Pada analitik perilaku pelanggan e-commerce, laporan per-kelas mengungkap variasi besar antara kelas dominan dan kelas yang jarang, sehingga akurasi agregat 90–95% tanpa presisi dan recall per-kelas berisiko menyesatkan interpretasi efektivitas nyata model di lapangan (Sunarya et al., 2024). Konteks riset kesehatan populasi pun memperlihatkan pola serupa: perbandingan model pada risiko penyebaran diabetes menekankan F1 dan recall sebagai pelengkap akurasi untuk menggambarkan keberimbangan deteksi di bawah ketidakseimbangan data

(Uhryn et al., 2025). Bahkan pada arsitektur modern seperti hibrida CNN–Transformer, evaluasi kinerja lazimnya dilaporkan bersamaan antara akurasi, F1, dan metrik lain untuk membaca keseimbangan generalisasi. Dalam diagnosis lesi oral (OSCC/leukoplakia), dampak klinis menuntut bacaan recall yang tinggi, sehingga evaluasi mempertimbangkan F1 dan presisi untuk memastikan deteksi sensitif tanpa lonjakan salah positif (Alam et al., 2024).

2.6.3. Presisi (Precision)

Presisi adalah ukuran yang digunakan untuk mengetahui tingkat ketepatan prediksi positif yang dihasilkan oleh model. Presisi menunjukkan seberapa banyak data yang diprediksi sebagai kelas positif benar-benar termasuk ke dalam kelas positif tersebut.

Rumus presisi adalah sebagai berikut:

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

Nilai presisi yang tinggi menunjukkan bahwa model memiliki tingkat kesalahan yang rendah dalam memberikan prediksi positif.

Presisi mengukur proporsi prediksi positif yang benar dari seluruh prediksi positif, sehingga menjadi kunci ketika salah positif mahal atau berdampak buruk. Recall (sensitivitas) mengukur proporsi kasus positif yang berhasil terdeteksi dari seluruh kasus positif, sehingga kritis pada skenario di mana missing positive tidak dapat ditoleransi.

2.6.4. F1-Score

F1-score adalah metrik evaluasi yang digunakan untuk menggabungkan nilai presisi dan recall dalam satu ukuran. F1-score dihitung sebagai rata-rata harmonis dari presisi dan recall. Metrik ini digunakan untuk memberikan gambaran yang lebih seimbang terhadap performa model, terutama ketika distribusi data antar kelas tidak seimbang.

Rumus F1-score adalah sebagai berikut:

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

F1-score digunakan untuk menilai keseimbangan antara kemampuan model dalam memberikan prediksi positif yang tepat dan kemampuannya dalam mendeteksi seluruh data positif. Semakin tinggi nilai F1-score, maka semakin baik performa model secara keseluruhan.

Keduanya saling tarik-menarik dan dibalanskan oleh F1-score, yang merupakan rata-rata harmonik antara presisi dan recall. Literatur menunjukkan penggunaan gabungan presisi, recall, dan F1 untuk menilai ketepatan sekaligus kelengkapan deteksi, terutama pada domain berisiko tinggi, dari keamanan penerbangan hingga kesehatan. Dalam penilaian model berbasis teks dan citra, pelaporan F1 turut menjadi standar karena mampu merangkum kinerja pada kondisi ketidakseimbangan kelas; misalnya konfigurasi hibrida RNN untuk klasifikasi teks tak terstruktur menunjukkan F1 tinggi sebagai indikasi keseimbangan yang baik antara salah positif dan salah negatif (Sunagar et al., 2024). Studi klasifikasi

material limbah dengan kerangka evaluasi deep learning juga memposisikan F1 berdampingan dengan akurasi untuk memastikan relevansi operasional di luar angka agregat (Al-Mashhadani, 2023).

Pada praktik uji banding model, AUC-ROC sering diikutsertakan guna menilai kemampuan diskriminasi model di berbagai ambang keputusan. AUC-ROC bersifat tidak peka ambang (threshold-insensitive) sehingga membantu membandingkan model secara lebih adil ketika kebijakan operasional dapat menggeser ambang prediksi. Beberapa studi lintas domain bertumpu pada paket metrik—akurasi, presisi, recall, F1, dan AUC-ROC—sebagai dasar pemilihan model yang siap diterapkan, alih-alih bertumpu pada satu angka (Lee, 2023). Pada penelitian e-commerce, laporan AUC bersanding dengan metrik per-kelas memperlihatkan bahwa dua model dengan akurasi serupa dapat berperilaku sangat berbeda saat ambang diubah; temuan semacam ini penting untuk kebijakan promosi atau stok yang sensitif terhadap kesalahan tipe tertentu (Sunarya et al., 2024).

Kesalahan umum dalam evaluasi model adalah hanya berfokus pada akurasi tanpa mempertimbangkan distribusi kelas yang tidak seimbang. Pada dataset penjualan Coffeeshop Bobskuy, perbedaan jumlah antara produk laris dan tidak laris menuntut penggunaan metrik lain seperti presisi, recall, dan F1-score untuk penilaian yang lebih objektif.

2.6.5. Recall

Recall adalah ukuran yang digunakan untuk mengetahui kemampuan model dalam menemukan seluruh data positif yang sebenarnya. Recall menunjukkan seberapa banyak data yang benar-benar termasuk kelas positif berhasil diprediksi

dengan benar oleh model.(Sukmana et al., 2024). Dalam penelitian ini, recall digunakan untuk melihat seberapa baik model mampu mendeteksi produk yang benar-benar Laris.

Rumus recall adalah sebagai berikut:

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

Nilai recall yang tinggi menunjukkan bahwa model mampu mengenali sebagian besar data positif dengan baik. Semakin tinggi nilai recall, maka semakin kecil kemungkinan model gagal mendeteksi produk yang sebenarnya termasuk dalam kategori Laris.

Pemodelan hepatitis menegaskan bahwa akurasi mesti dibaca bersama presisi, recall, dan F1 untuk menilai utilitas klinis (Chiawa & Chibueze, 2025). Di sisi lain, adanya skor mendekati sempurna—bahkan 100%—pada dataset benchmark tertentu setelah penalaan bayesian mengisyaratkan risiko overfitting atau ketidakgeneralisasian; ini menggarisbawahi pentingnya validasi ketat dan evaluasi pada data yang berbeda domain. Karena itu, pelaporan per-kelas dan matriks kebingungan menjadi alat diagnosis yang vital untuk membuka di mana model unggul dan di mana ia gagal, yang selama ini dibuktikan di studi klasifikasi keluhan publik berbasis teks dan perbandingan KNN, NB, serta C4.5 (Bahri et al., 2022).

Selain angka kinerja, isu reliabilitas dan keterjelasan (explainability) model makin dipandang sebagai bagian dari evaluasi yang bertanggung jawab. Pada domain kesehatan mental, adopsi *Explainable AI (XAI)* digunakan untuk menilai keandalan model, sehingga pemegang keputusan tidak bergantung semata pada metrik numerik namun juga pada alasan prediksi yang dapat diaudit (Rafieinasab et al., 2024). Perspektif MLOps menambah dimensi temporal terhadap evaluasi: performa harus dimonitor setelah deployment untuk menangkap drift data dan degradasi kinerja, tidak cukup berhenti pada laporan metrik di set uji statis (Bodor et al., 2023). Dalam konteks operasional seperti ritel dan e-commerce, pemantauan berkelanjutan atas akurasi, presisi, recall, F1, serta alarm jika metrik turun di bawah ambang merupakan praktik yang disarankan untuk menjaga mutu prediksi.

Adopsi protokol evaluasi multi-metrik mendapat dukungan kerangka skoring agregat untuk mempermudah perbandingan model lintas metrik. Gagasan *Machine Learning Cumulative Performance Score (MLCPS)* menawarkan cara menghimpun beberapa metrik menjadi skor komposit yang dapat mengurangi bias menuju satu metrik tertentu, sembari tetap transparan terhadap komponen pembentuknya (Akshay et al., 2022). Walau skor komposit tidak menggantikan pembacaan granular presisi/recall per-kelas, ia berguna saat menyaring banyak kandidat model ke dalam shortlist berbasis bukti yang konsisten dan dapat diaudit.

Ketika menilai *Naïve Bayes* pada prediksi kelarisan, pendekatan yang seimbang meliputi: pelaporan akurasi sebagai gambaran umum; inspeksi presisi dan recall per-kelas untuk memahami biaya salah klasifikasi pada produk laris vs tidak laris; serta penggunaan F1 guna menyeimbangkan kebutuhan menekan false

positive dan false negative. Jika keputusan bisnis mengutamakan deteksi produk berpotensi laris (positif) semaksimal mungkin, recall perlu ditargetkan tinggi; namun bila salah positif berdampak pada penumpukan stok, presisi perlu dikendalikan. Kasus-kasus lintas domain memperlihatkan bagaimana F1 yang kuat sering kali beriringan dengan recall yang memadai saat biaya kehilangan positif tinggi.

Evaluasi juga harus mempertimbangkan realisme eksperimental. Skor yang sangat tinggi pada dataset kecil atau setelah penalaan intensif perlu diverifikasi dengan uji silang yang ketat dan set uji eksternal. Jebakan data leakage, overfitting pada benchmark spesifik, atau distribusi uji yang terlalu mirip dengan latih dapat menimbulkan optimisme palsu, sebagaimana diperingatkan pada studi penalaan berbasis bayesian di domain kanker payudara (Alsabry & Algabri, 2024). Oleh karena itu, pelaporan kinerja sebaiknya disertai prosedur validasi yang jelas, representasi matriks kebingungan, dan—bila mungkin—hasil pada data berbeda domain untuk menilai portabilitas model. Dalam proyek teks real-world, misalnya, klasifikasi keluhan publik menunjukkan betapa laporan per-kelas dan matriks kebingungan membantu menginformasikan keputusan rekayasa fitur lanjutan dan penyesuaian ambang (Bahri et al., 2022). Demikian pula, komparatif model di kesehatan—deteksi penyakit ginjal dan prediksi hepatitis—memperlihatkan bahwa pemilihan model bukan sekadar angka tertinggi, melainkan keseimbangan metrik yang sesuai dengan tujuan klinis (Uddin et al., 2024) ;(Chiawa & Chibueze, 2025).

Terdapat juga penerapan alur praktik untuk Coffeeshop Bobskuy dengan *Naïve Bayes*: pertama, tetapkan tujuan operasional yang menerjemahkan risiko

salah positif/negatif menjadi preferensi metrik; kedua, laporkan akurasi, presisi, recall, F1, serta metrik per-kelas dan matriks kebingungan; ketiga, gunakan skema agregasi seperti MLCPS hanya sebagai alat bantu komparatif, bukan pengganti analisis rinci; keempat, terapkan validasi silang dan, jika memungkinkan, uji eksternal; kelima, rangkai evaluasi dengan kerangka XAI untuk meningkatkan kepercayaan pemangku kepentingan; dan terakhir, integrasikan pemantauan metrik dalam pipeline MLOps untuk memastikan kinerja tetap terjaga setelah model dideploy. Dengan demikian, evaluasi tidak berhenti pada laporan satu kali, melainkan menjadi proses berkelanjutan yang menyeimbangkan angka, konteks, dan keandalan.

Keseluruhan bukti literatur yang digunakan memperkuat posisi bahwa akurasi dan presisi merupakan pondasi, tetapi bukan kesimpulan akhir dalam menilai efektivitas model. Recall dan F1 mengisi celah penting pada skenario ketidakseimbangan dan biaya asimetris, sementara AUC-ROC, laporan per-kelas, matriks kebingungan, serta kerangka XAI dan MLOps memastikan bahwa evaluasi bersifat menyeluruh, dapat dipercaya, dan berkelanjutan. Kerangka penilaian demikian selaras dengan kebutuhan praktis bisnis ritel seperti Coffeeshop Bobskuy, yang menuntut keputusan cepat tetapi tetap akuntabel terhadap risiko operasional—dan literatur lintas domain menyediakan bukti yang cukup untuk menjadikan pendekatan multi-metrik sebagai standar evaluasi yang bertanggung jawab (Sukmana et al., 2024);(Pendyala & Kim, 2024).

2.7. Alat Bantu Pemrograman dan Tools Pendukung



Gambar 2. 2 Software Orange

Pemilihan Orange Data Mining sebagai alat bantu utama dalam penelitian ini didasarkan pada kemampuannya menyediakan workflow visual yang memudahkan proses implementasi machine learning, terutama bagi pengguna yang tidak memiliki latar belakang pemrograman. Orange menyediakan lingkungan berbasis widget untuk membangun pipeline analisis data mulai dari input dataset, preprocessing, pemodelan, hingga visualisasi. Platform ini mendukung integrasi dengan format dataset yang umum seperti CSV atau Excel, yang banyak digunakan dalam penelitian bisnis dan analisis penjualan. Orange memungkinkan pembuatan model Naïve Bayes secara intuitif tanpa perlu menulis kode, yang membuatnya cocok digunakan dalam penelitian ini yang berfokus pada analisis penjualan produk di Coffeeshop Bobskuy.

2.7.1. Praproses Data

Langkah pertama dalam workflow Orange adalah praproses data, yang mencakup pembersihan data dari nilai kosong (missing values), penghapusan duplikasi, dan normalisasi. Untuk itu, Orange menyediakan berbagai widget, seperti

Select Columns untuk memilih atribut input dan target, serta *Edit Domain* yang memungkinkan penyesuaian terhadap atribut yang digunakan dalam pemodelan. Praproses ini juga mencakup penyesuaian atau transformasi data agar siap digunakan oleh algoritma *Naïve Bayes*. Jika data memiliki komponen teks, Orange memiliki modul *Corpus* dan *TF-IDF* untuk mengonversi data teks menjadi representasi numerik yang sesuai untuk analisis.

2.7.2. Ekstraksi Fitur

Orange menyediakan widget *Corpus* dan *TF-IDF* untuk menangani data teks, yang memungkinkan representasi data dalam bentuk numerik yang sesuai untuk algoritma *Naïve Bayes* (Sunarya et al., 2024). Representasi fitur yang akurat memastikan bahwa probabilitas dalam algoritma *Naïve Bayes* dapat dihitung dengan benar.

2.7.3. Implementasi Model *Naïve Bayes*

Dalam Orange, pengguna hanya perlu menghubungkan widget *Naïve Bayes* dengan *Test & Score* untuk melatih dan menguji model pada dataset yang telah disiapkan. Proses ini dapat dilakukan sepenuhnya secara modular melalui antarmuka grafis. Orange memungkinkan pengguna untuk melakukan eksperimen dengan parameter default maupun variasi distribusi (*Gaussian* atau *Multinomial*) tanpa memerlukan skrip kode manual. Keunggulan utama Orange adalah kemudahan dalam mencoba berbagai konfigurasi dan langsung melihat hasil evaluasi model.

2.7.4. Evaluasi Performa Model

Setelah model *Naïve Bayes* diterapkan, evaluasi performa model dilakukan dengan menggunakan *widget Test & Score* dan *Confusion Matrix* di Orange. Metrik evaluasi yang digunakan antara lain akurasi, presisi, recall, dan F1-score, yang penting untuk menilai performa model dalam mengklasifikasikan produk sebagai “laris” atau “tidak laris”. *Confusion Matrix* akan menunjukkan perbandingan antara prediksi dan label yang sebenarnya, sementara metrik lainnya membantu dalam mengukur kualitas prediksi. Evaluasi ini sangat penting untuk memastikan bahwa model *Naïve Bayes* dapat menghasilkan prediksi yang akurat dan dapat diandalkan.

2.7.5. Visualisasi

Visualisasi merupakan bagian integral dari pemahaman hasil klasifikasi dan analisis data. Orange menyediakan berbagai widget visual seperti *Scatter Plot*, *Box Plot*, dan *ROC Analysis* untuk menampilkan performa model *Naïve Bayes* secara intuitif. Melalui visualisasi yang jelas, pemangku kepentingan dapat melihat hasil yang telah dievaluasi dan mengambil keputusan yang lebih baik terkait stok dan promosi produk.

Untuk memastikan bahwa model *Naïve Bayes* tidak terpengaruh oleh bias evaluasi, digunakan teknik *k-fold cross-validation*. *Widget Test & Score* di Orange memungkinkan pengguna untuk melakukan validasi silang otomatis, membagi dataset menjadi beberapa subset yang digunakan untuk pelatihan dan pengujian model secara bergantian. Teknik ini memastikan bahwa performa model tidak hanya baik pada data pelatihan, tetapi juga pada data yang belum pernah dilihat sebelumnya, sehingga model dapat digeneralisasi dengan baik terhadap data baru.

2.7.6. Interpretasi Hasil

Langkah terakhir dalam proses ini adalah menginterpretasikan hasil klasifikasi dan menggunakannya untuk pengambilan keputusan operasional. Output probabilistik dari model *Naïve Bayes* memungkinkan peneliti untuk merumuskan rekomendasi berbasis peluang, seperti kemungkinan suatu produk menjadi laris berdasarkan pola penjualan historis. Hasil evaluasi dapat diekspor dalam format Excel atau laporan interaktif untuk dianalisis lebih lanjut, yang dapat digunakan oleh manajemen Coffeeshop Bobskuy untuk merencanakan stok produk, memberikan diskon, atau merancang strategi promosi yang lebih efektif.

Orange mendukung penyimpanan workflow dalam format file .ows, yang memungkinkan pengguna untuk membuka dan mengedit proyek yang sama di masa depan. Dengan dokumentasi yang baik, model *Naïve Bayes* yang dikembangkan dapat diperbarui dan disesuaikan dengan data baru tanpa kehilangan rekam jejak eksperimen.

2.8. Penelitian Terdahulu

Tabel 2. 1 Penelitian terdahulu

Peneliti	judul	Tahun	Data dan Algoritma	Hasil
Djami et al.	Prediksi tingkat penjualan produk di PT Kotamas Bali	2023	Data penjualan PT Kotamas Bali; Naïve Bayes (baseline, dibanding KNN); penekanan pra-proses	Akurasi 63,414%; kinerja sangat bergantung kualitas data & pra-proses
Zahri et al.	Analisis penjualan produk dengan klasifikasi Naïve Bayes	2024	Data penjualan ritel; Naïve Bayes dengan fitur yang relatif independen	Akurasi 83,3%; Presisi 84,2%; Recall 88,9%; efektif untuk prediksi produk

				laris & dukung stok/promosi
Dewi et al.	Prediksi repeat order & kategori jumlah penjualan	2024	Data pelanggan; perbandingan NB vs Decision Tree vs Random Forest	NB 54% (kategori jumlah penjualan); 92% (klasifikasi distribusi lokasi pelanggan)
Harahap et al.	Prediksi pembelian cat dari data pemasok	2021	60 sampel pembelian; Naïve Bayes	Akurasi 80%; stabil pada kategori produk tertentu dengan data terbatas
Suryani et al.	Analisis penjualan kaca & perencanaan inventori	2023	Data penjualan/inventori; NB dalam skema hibrida; dibanding GRNN	NB kompetitif meski bukan model utama; relevan untuk sistem hibrida
Penelitian ini	Implementasi Algoritma Naïve Bayes untuk Klasifikasi Penjualan Produk pada Coffeeshop Bobskuy	2025	Data transaksi penjualan Coffeeshop; visualisasi; Orange	Sedang berjalan

Kebaruan dalam penerapan algoritma *Naïve Bayes* pada konteks penjualan Coffeeshop Bobskuy dengan mengintegrasikan analisis klasifikasi berbasis data transaksi aktual yang diproses dan divisualisasikan menggunakan platform Orange. Berbeda dari penelitian sebelumnya yang umumnya berfokus pada prediksi

penjualan dalam skala ritel besar atau studi komparatif antar-algoritma, penelitian ini menekankan penerapan langsung pada bisnis kecil-menengah dengan karakteristik produk, pola pembelian, dan fluktuasi permintaan yang lebih variatif. Kebaruan lainnya terletak pada penggunaan kombinasi pra-pemrosesan khusus untuk data makanan dan minuman, penyusunan fitur yang disesuaikan dengan perilaku konsumen harian, serta fokus pada prediksi kategori produk laris yang dapat digunakan secara langsung untuk pengambilan keputusan operasional.